

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 20:50:10 : Search time 465 Seconds

(Without alignments)
10417.308 Million cell updates/sec

Title: US-09-601-168b-1

Perfect score: 2151

Sequence: 1 tgcgttgccgcgcgcctggc.....gtttgcagagaaaaaaa 2151

Scoring table: IDENTITY-NUC

Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_101002:*

- 1: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT.*
- 2: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT.*
- 3: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT.*
- 4: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT.*
- 5: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT.*
- 6: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT.*
- 7: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT.*
- 8: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT.*
- 9: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT.*
- 10: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT.*
- 11: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT.*
- 12: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT.*
- 13: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT.*
- 14: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT.*
- 15: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT.*
- 16: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT.*
- 17: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT.*
- 18: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT.*
- 19: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT.*
- 20: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT.*
- 21: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT.*
- 22: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT.*
- 23: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT.*
- 24: /SID2/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2151	100.0	2151	20	AAK86501
2	2151	100.0	2151	21	AAK51229
3	2151	100.0	2151	21	AAK93350
4	2151	100.0	2151	21	AAK93710
5	2151	100.0	2151	22	AAK84610
6	2120.8	98.6	3220	22	AAH90079
7	2114.8	98.3	2419	21	AAK29233
8	2011.4	93.5	3622	22	AAH89966
9	2008.2	93.4	3003	22	AAK52699

10	2008.2	93.4	3003	22	AAK52700	Human polynucleoti
11	2008.2	93.4	3003	21	AAK52701	Human polynucleoti
12	1707	79.4	1707	21	AAK73132	Human beta-transdu
13	1697.8	78.9	2285	22	AAK51715	Human polynucleoti
14	1697.8	78.9	2366	22	AAK51717	Human polynucleoti
15	1688.8	78.5	2207	22	AAK51716	Human polynucleoti
16	1682.2	78.2	2175	21	AAK93714	F-box protein FWD1
17	1443	67.1	1707	21	AAK73131	Mouse ubiquitin 11
18	883.4	41.1	2084	22	AAK61150	Human E3 ubiquitin
19	883.4	41.1	4230	21	AAK51228	Human polynucleoti
20	883.4	41.1	4344	22	AAK52260	Human polynucleoti
21	883.4	41.1	4360	22	AAK59364	Human polynucleoti
22	866.4	31.4	2387	23	ABK06081	Drosophila melanog
23	657	30.5	657	24	ABK86903	Human beta trcp (b
24	572.6	26.6	647	20	AAK09025	Differentiation in
25	518.8	24.1	1561	22	AAK84601	Beta-TRCP/N/SKP2.C
26	435.4	20.2	951	23	ABK43155	CDNA encoding huma
27	387.6	18.0	594	23	ABV58856	Human prostate exp
28	288	13.4	9129	23	ABK06080	Drosophila melanog
29	272	12.6	479	22	ABK44075	Human breast cell
30	272	12.6	479	22	ABK44075	Human breast cell
31	272	12.6	479	22	ABK44075	Human breast cell
32	272	12.6	479	22	ABK44075	Human breast cell
33	272	12.6	479	22	ABK44075	Human breast cell
34	272	12.6	479	22	ABK44075	Human breast cell
35	272	12.6	479	22	ABK44075	Human breast cell
36	272	12.6	479	22	ABK44075	Human breast cell
37	272	12.6	479	22	ABK44075	Human breast cell
38	208.4	9.7	370	22	ABK44075	Human breast cell
39	208.4	9.7	370	22	ABK44075	Human breast cell
40	208.4	9.7	370	22	ABK44075	Human breast cell
41	208.4	9.7	370	22	ABK44075	Human breast cell
42	208.4	9.7	370	22	ABK44075	Human breast cell
43	208.4	9.7	370	22	ABK44075	Human breast cell
44	208.4	9.7	370	22	ABK44075	Human breast cell
45	208.4	9.7	370	22	ABK44075	Human breast cell

ALIGNMENTS

RESULT 1
AAK86501
ID AAK86501 standard; CDNA: 2151 BP.

AC AAK86501;
XX 30-SEP-1999 (first entry)

DE CDNA encoding a human beta-transducin repeat containing protein.

XX Beta-transducin repeat containing protein; beta-trcp; SKP1;
XX Proteosome degradation pathway; Vpu protein; beta-catenin;
XX human immune deficiency virus-1; HIV-1; cellular protein; IKAPAB;
XX ubiquitinylation; phosphorylated protein; tumour; apoptosis; Alzheimer's;
XX antiviral; antitumour; cell cycle regulation; protein degradation;
XX anti-inflammation; osteo articular inflammation; acute inflammation;
XX tumour necrosis factor; ss.

OS Homo sapiens.

XX Location/Qualifiers

XX Key 70..1779
XX CDS /tag= a
XX /product= beta-Trcp
XX /note= "beta-transducin repeat containing protein"

XX MO9938969-A1.

XX 05-AUG-1999.

XX 29-JAN-1999; 99WO-FR00196.

PR	09-DEC1998:	98FR-0015545.	
PR	30-JAN-1998:	98FR-0001100.	
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.		
PA	(INSP) INSE PASTEUR.		
PI	Arenzana Seisdedos F, Benarous R, Concordet J, Durand H;		
PI	Kroll M, Margottin F;		
XX	WPI; 1999-469329/39.		
DR	P-PSDB; AAY24054.		
XX			
PT	New human beta-transducin repeat containing protein and its		
PT	fragments useful as, or to screen for, antiviral, antitumour,		
PT	anti-inflammatory and anti-Alzheimer's agents		
PS	Claim 7; Page 57-60; 71pp; French.		
XX			
CC	The present sequence encodes a human beta-transducin repeat containing		
CC	protein (beta-Trcp). The protein directs proteins to the proteosome		
CC	degradation pathways. The protein is able to interact with the Vpu		
CC	protein of human immune deficiency virus-1 (HIV-1), cellular proteins		
CC	IkappaB or beta-catenin (bc) and/or protein Skp1. The protein controls		
CC	ubiquitinylation of phosphorylated proteins and thus their targeting to		
CC	proteosomes for degradation. Depending on whether the process is		
CC	inhibited or promoted, the result may be delayed breakdown of CD4 (in		
CC	cases of HIV-1 infection); increased activity of IkB (and thus reduced		
CC	activity of NFkappaB) and increased degradation of mutant bc in tumour		
CC	cells, or increased bc survival (and reduced apoptosis) in Alzheimer's		
CC	nucleus. The beta-Trcp protein, and its active peptide fragments, or its		
CC	nucleic acid, are used to screen for anti HIV-1 agents (antiviral),		
CC	antitumour agents that disrupt cell cycle regulation or protein		
CC	degradation in human tumour cells, and anti-inflammatory agents that		
CC	disrupt activation by NFkappaB. Fragments of the protein are also		
CC	useful for treating osteo-articular inflammation or acute inflammation		
CC	associated with release of tumour necrosis factor.		
XX			
XX			
SO	Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;		
	Query Match 100.0%; Score 2151; DB 20; Length 2151;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 TCGGTTGGCTGGGCGCTGGCACCAAGGGGCGGCCCGGCGAGAGCGGACCTGGGCC 60		
DB	1 TCGGTTGGCTGGGCGCTGGCACCAAGGGGCGGCCCGGCGAGAGCGGACCTGGGCC 60		
QY	61 TCGGCGATTATGAGCCCGGCGAGGCGGTGCTGCGCAAGGAAGGCGCTCAAGTTTGTGAT 120		
DB	61 TCGGCGATTATGAGCCCGGCGAGGCGGTGCTGCGCAAGGAAGGCGCTCAAGTTTGTGAT 120		
QY	121 TCCTAGAGAGAGAAGAGACTGTATATATATGCGCAACCCCTAGGAAGATTAATACAGAGAAG 180		
DB	121 TCCTAGAGAGAGAAGAGACTGTATATATGCGCAACCCCTAGGAAGATTAATACAGAGAAG 180		
QY	181 AATTCACCTTAGACAGACATACACAGCGTGGCCAGACTCTCTTAACCAAGAAACAGTA 240		
DB	181 AATTCACCTTAGACAGACATACACAGCGTGGCCAGACTCTCTTAACCAAGAAACAGTA 240		
QY	241 TGTTTAGCAAGCACTGCTATGAAGACTAGAAATGTGTGGCGAAACAAACCTTGCCAT 300		
DB	241 TGTTTAGCAAGCACTGCTATGAAGACTAGAAATGTGTGGCGAAACAAACCTTGCCAT 300		
QY	301 GGCACCTTCACATATGATTTGGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360		
DB	301 GGCACCTTCACATATGATTTGGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360		
QY	361 AAGGAACGTGTGTCAATAATCTTGAGAGAGTGTGCAAGTCAAGTCAAGGAGATTGTG 420		
DB	361 AAGGAACGTGTGTCAATAATCTTGAGAGAGTGTGCAAGTCAAGTCAAGGAGATTGTG 420		
QY	421 GAAATCTTATATCCCAATGTGCAATTACCAACATGGGCAATMAACTGTATCTTAA 480		
DB	421 GAAATCTTATATCCCAATGTGCAATTACCAACATGGGCAATMAACTGTATCTTAA 480		

Db	421	GAACATCTTATATCCAAATGIGTATTTACCAACATGGGACATTAACCTGATCTTTAA	480
Qy	481	CCATGTTGCAGAGAGATTTTCATTAACCTGCTGCCAGCTCGGGAGTTGCATATTCGT	540
Db	481	CCATGTTGCAGAGAGATTTTCATTAACCTGCTGCCAGCTCGGGAGTTGCATATTCGT	540
Qy	541	GAGAACATCTGTGATACCTGGATGGATGCCAAATACATATGCTGCTGAACCTTGTCAG	600
Db	541	GAGAACATCTGTGATACCTGGATGGATGCCAAATACATATGCTGCTGAACCTTGTCAG	600
Qy	601	GAATGTAACCGAGTGAACCTCTGATGAGCATTCCTGTGGAGAAAGCTTATACGAGAAATGATC	660
Db	601	GAATGTAACCGAGTGAACCTCTGATGAGCATTCCTGTGGAGAAAGCTTATACGAGAAATGATC	660
Qy	661	AGGACAGATTCCTGTGGAGAGCCCTGGCACAAGAGAGATGGGACATTTATATTC	720
Db	661	AGGACAGATTCCTGTGGAGAGCCCTGGCACAAGAGAGATGGGACATTTATATTC	720
Qy	721	AAAAAACAACCTCTGCACGGGAAATGCTCTCCCAACTTTTATATAGACACTTATTCCT	780
Db	721	AAAAAACAACCTCTGCACGGGAAATGCTCTCCCAACTTTTATATAGACACTTATTCCT	780
Qy	781	AAAAATATACAGACATTTGAGACATTAATATCTAATTGAGATGTGGAAACATATGTTTA	840
Db	781	AAAAATATACAGACATTTGAGACATTAATATCTAATTGAGATGTGGAAACATATGTTTA	840
Qy	841	CAGAGAAATTCAGTGGCGGAAGTGAACCAAGCAAGAGATTTACGTTTACAGTATGTGAT	900
Db	841	CAGAGAAATTCAGTGGCGGAAGTGAACCAAGCAAGAGATTTACGTTTACAGTATGTGAT	900
Qy	901	CAGAAATATGTAGGGGCTTCGAGACACACATCAATCAAGTCTGGATTAACCAACATTTG	960
Db	901	CAGAAATATGTAGGGGCTTCGAGACACACATCAATCAAGTCTGGATTAACCAACATTTG	960
Qy	961	GAATGCACGCAATTCCTCACAGGCATACAGGTTCACTCCTCTGTCCAGATATGATAG	1020
Db	961	GAATGCACGCAATTCCTCACAGGCATACAGGTTCACTCCTCTGTCCAGATATGATAG	1020
Qy	1021	AGAGTATCATTAACAGGATCTCGGATTCACAGCTCAGACTGTGGATTAATACAGGT	1080
Db	1021	AGAGTATCATTAACAGGATCTCGGATTCACAGCTCAGACTGTGGATTAATACAGGT	1080
Qy	1081	GAATGCTAAACACGTTGATTCACCATTTGGAAGAGCTTCGTGACTGTGGTTCAATAT	1140
Db	1081	GAATGCTAAACACGTTGATTCACCATTTGGAAGAGCTTCGTGACTGTGGTTCAATAT	1140
Qy	1141	GGCATGATGTGACCTGCTCCAAAGATCGTTCCATTTGCTGTATGGATATGGCTCCCCA	1200
Db	1141	GGCATGATGTGACCTGCTCCAAAGATCGTTCCATTTGCTGTATGGATATGGCTCCCCA	1200
Qy	1201	ACTGCATTTACCCCTCGGAGGCTGCTGGTGGGACACCGAGCTGTGCAATGTTGTAGAC	1260
Db	1201	ACTGCATTTACCCCTCGGAGGCTGCTGGTGGGACACCGAGCTGTGCAATGTTGTAGAC	1260
Qy	1261	TTTGTATGACAGTACATTTGTTTCGATCTGGGGATATAGAACTATAAGATATGACACA	1320
Db	1261	TTTGTATGACAGTACATTTGTTTCGATCTGGGGATATAGAACTATAAGATATGACACA	1320
Qy	1321	AGTACTGTGAAATTTGTAAGGACCTTAATGACACAAACGAGCAATTCCTGTTTCCAG	1380
Db	1321	AGTACTGTGAAATTTGTAAGGACCTTAATGACACAAACGAGCAATTCCTGTTTCCAG	1380
Qy	1381	TACAGGCAACAGCTGATGAGTGTGCTCATGTCGACAAACATATCAGATTAATGGACATA	1440
Db	1381	TACAGGCAACAGCTGATGAGTGTGCTCATGTCGACAAACATATCAGATTAATGGACATA	1440
Qy	1441	GAATGTGCTGATGTTTACGATGTTTGAAGAGCCATGAGGAATTTGTCCTGTTATTCGA	1500
Db	1441	GAATGTGCTGATGTTTACGATGTTTGAAGAGCCATGAGGAATTTGTCCTGTTATTCGA	1500
Qy	1501	TTTATATTAACAAGAGATAGTACAGTGGGCGCTTATGATGAAAAATTAAGTGTGGGATCTT	1560
Db	1501	TTTATATTAACAAGAGATAGTACAGTGGGCGCTTATGATGAAAAATTAAGTGTGGGATCTT	1560

```

QY 1561 GTGGCTGTTTGGAGACCCCGCTCTCTGACAGAGACACTCTCTTACGAGACCCCTTGTGAG 1620
    |||||||
DB 1561 GTGGCTGTTTGGAGACCCCGCTCTCTGACAGAGACACTCTCTTACGAGACCCCTTGTGAG 1620
QY 1621 CATTCGCGAAGATTTTTCGACTACAGTTTATGATATTCAGATGTCAGTTCACAT 1680
    |||||||
DB 1621 CATTCGCGAAGATTTTTCGACTACAGTTTATGATATTCAGATGTCAGTTCACAT 1680
QY 1661 GATGACACAAATCTCATCTGGGACTCTCTTAATGATCCAGCTGCCAAGCTGAACCCCCC 1740
    |||||||
DB 1661 GATGACACAAATCTCATCTGGGACTCTCTTAATGATCCAGCTGCCAAGCTGAACCCCCC 1740
QY 1741 CGTTCCTCTTCTGACACATACCTACATCTCCGATTAATTAACCATACACTGACTCAT 1800
    |||||||
DB 1741 CGTTCCTCTTCTGACACATACCTACATCTCCGATTAATTAACCATACACTGACTCAT 1800
QY 1801 ACTTGCCGAGAGACCCATTAAGTTGGGTTATTTAACTGATCGCAATACAGATGAGC 1860
    |||||||
DB 1801 ACTTGCCGAGAGACCCATTAAGTTGGGTTATTTAACTGATCGCAATACAGATGAGC 1860
QY 1861 AACCACTAACAATCAAACTACTGCCCCAGTTTCCCTGGAGTACCGAGAGAGAGGCTT 1920
    |||||||
DB 1861 AACCACTAACAATCAAACTACTGCCCCAGTTTCCCTGGAGTACCGAGAGAGAGGCTT 1920
QY 1921 TGAGACTCTGTTGGAGACAGTTGCTGTGACGTCCGCGCCAGAGCGTCTACTAGACACA 1980
    |||||||
DB 1921 TGAGACTCTGTTGGAGACAGTTGCTGTGACGTCCGCGCCAGAGCGTCTACTAGACACA 1980
QY 1981 ACTGACTCTTCACTGCTGCTATCAGAAAGATGCTCTTCAATTTGAATGATTTGGAAC 2040
    |||||||
DB 1981 ACTGACTCTTCACTGCTGCTATCAGAAAGATGCTCTTCAATTTGAATGATTTGGAAC 2040
QY 2041 TTTTAACTCCCTCCTCTCTCTCTTACCTGTGACAGCTAGTTTTCCTCATTTGCTTC 2100
    |||||||
DB 2041 TTTTAACTCCCTCCTCTCTCTCTTACCTGTGACAGCTAGTTTTCCTCATTTGCTTC 2100
QY 2101 CAGACAAAGTGACTATTAATATTTAGTGTGGCAGAAAAA 2151
    |||||||
DB 2101 CAGACAAAGTGACTATTAATATTTAGTGTGGCAGAAAAA 2151

RESULT 2
ID AAA51229 standard; DNA: 2151 BP.
AC AAA51229;
XX 26-SEP-2000 (first entry)
DT Human beta-TrCP coding sequence.
DE
XX E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
KW anti-inflammatory; immunosuppressive; cytosolic; ds.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 70..1779
FT CDS /*tag= a
FT /product= Human_TrCP
XX
XX WO200034447-A2.
XX
XX 15-JUN-2000.
XX
XX 10-DEC-1999; 99WO-US29371.
XX
XX 10-DEC-1998; 98US-021060.
XX
XX (SIGN-) SIGNAL PHARM INC.
XX (YISS) YISSON RES & DEV CO.

```

```

XX Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzudai A;
PI Lavon I, Yaron A;
XX
XX WPI: 2000-431294/37.
DR P-PDB; AAY96697.
XX
PT Polypeptide enhancing phosphorylated Ikkappa ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX
PS Disclosure: Page 72; 77pp; English.
XX
XX This DNA encodes human beta-TrCP, an F-box/WD protein family member,
XX which has been shown to have homology to human E3 ubiquitin ligase (E3).
XX E3 enhances ubiquitination of phosphorylated I-kappa-B, an inhibitor
XX of nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
XX degradation via the ubiquitin pathway is useful for identifying
XX modulators of this process for use in treating diseases associated with
XX activation of NF-kappa-B. In vitro analysis suggests that deletion of
XX the F-box results in a protein that functions as a dominant negative
XX molecule in vivo. Transient over-expression of delta-beta-TrCP (a
XX deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
XX in stimulated Jurkat cells, resulting in accumulation of phosphorylated
XX I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
XX activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
XX inflammatory diseases, autoimmune diseases, cancer and viral infections.
XX
SO Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 21; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGTTGGCTGGCGCCGCGACCAAGGGGCGCCCGGGGAGAGCGGACCCAGTGGCC 60
    |||||||
DB 1 TGGGTTGGCTGGCGCCGCGACCAAGGGGCGCCCGGGGAGAGCGGACCCAGTGGCC 60
QY 61 TGGGCGATTATGAGACCGCGCGGAGGGGCTGCTGCAAGAGAGGCACTCAAGTTTAAAT 120
    |||||||
DB 61 TGGGCGATTATGAGACCGCGCGGAGGGGCTGCTGCAAGAGAGGCACTCAAGTTTAAAT 120
QY 121 TCCTCAGAGAGAGAAAGACTTAATATGGGGAACCCCTTAGAGAAATATACAGAGAAG 180
    |||||||
DB 121 TCCTCAGAGAGAGAAAGACTTAATATGGGGAACCCCTTAGAGAAATATACAGAGAAG 180
QY 181 AATTCACTTAGACAGACATCAACAGCTGTGCCAGACTGCTTAACCAAGAAACAGTA 240
    |||||||
DB 181 AATTCACTTAGACAGACATCAACAGCTGTGCCAGACTGCTTAACCAAGAAACAGTA 240
QY 241 TGTTCAGACAGACGTGTATGAAGACTGAGAAATTTGTGGCCAAACAAACTTGGCAAT 300
    |||||||
DB 241 TGTTCAGACAGACGTGTATGAAGACTGAGAAATTTGTGGCCAAACAAACTTGGCAAT 300
QY 301 GGCACCTTCCAGTATGATTTGGCCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAAAGAA 360
    |||||||
DB 301 GGCACCTTCCAGTATGATTTGGCCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAAAGAA 360
QY 361 AAGGAAGTGTGTCAAAATCTTTGAGCAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGG 420
    |||||||
DB 361 AAGGAAGTGTGTCAAAATCTTTGAGCAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGG 420
QY 421 GAACATCTTATCCCAATGCTCTTACCAACATGGGACATTAACCTGATCTTAA 480
    |||||||
DB 421 GAACATCTTATCCCAATGCTCTTACCAACATGGGACATTAACCTGATCTTAA 480
QY 481 CCTATGTCAGAGAGATTTCACTGCTCTGCCAGCTCGGGGATGGATCATATCGCT 540
    |||||||
DB 481 CCTATGTCAGAGAGATTTCACTGCTCTGCCAGCTCGGGGATGGATCATATCGCT 540
QY 541 GAGAACCTTGTCTACATCCGATGCGCAATCACTATGCTGCTGCACTGTCGCAAG 600
    |||||||

```

Db 541 GAGACATTTCTGTCAATACCTGATGCCAAATCACTATGTCTGTCGACTGTGTGCAAG 600
 QY 601 GAATGTACCGAGTGAACCTGTGATGGCATGTGTGGAGAAGCTTATTCAGAGATGTGTC 660
 Db 601 GAATGTACCGAGTGAACCTGTGATGGCATGTGTGGAGAAGCTTATTCAGAGATGTGTC 660
 QY 661 AGCAGAGATTTCTGTGTGGAGAGGCTTGCGACAAAGAGAGATGGGAGACATATTTATTC 720
 Db 661 AGCAGAGATTTCTGTGTGGAGAGGCTTGCGACAAAGAGAGATGGGAGACATATTTATTC 720
 QY 721 AAAAACAACCTCTGCGAGGGAATGCTCCCTCCCACTCTTTTATGTAGACACTTTATCTCT 780
 Db 721 AAAAACAACCTCTGCGAGGGAATGCTCCCTCCCACTCTTTTATGTAGACACTTTATCTCT 780
 QY 781 AAAATATACAGACATTTGAGACATATGATTTGAGATGTGGAAGACATATTTTA 840
 Db 781 AAAATATACAGACATTTGAGACATATGATTTGAGATGTGGAAGACATATTTTA 840
 QY 841 CAGAGATTTCTGTGTGGAGAGGCTTGCGACAAAGAGAGATGGGAGACATATTTATTC 900
 Db 841 CAGAGATTTCTGTGTGGAGAGGCTTGCGACAAAGAGAGATGGGAGACATATTTATTC 900
 QY 901 CAGAAATATGTAAGCGGCTTGCGACAAACACATCAAGATCGGGATAAAAACACATTG 960
 Db 901 CAGAAATATGTAAGCGGCTTGCGACAAACACATCAAGATCGGGATAAAAACACATTG 960
 QY 961 GAATGCAAGCAATTTCTCAAGAGGCTTACAGATTCAGTCTGCTGTCCAGATATGATGAG 1020
 Db 961 GAATGCAAGCAATTTCTCAAGAGGCTTACAGATTCAGTCTGCTGTCCAGATATGATGAG 1020
 QY 1021 AGAGTGTCTATACAGAGATTCAGATTCAGGCTGAGAGATGGGAGATGTAATACAGT 1080
 Db 1021 AGAGTGTCTATACAGAGATTCAGATTCAGGCTGAGAGATGGGAGATGTAATACAGT 1080
 QY 1081 GAAATGCTAAACAGGTTGATTCACATTTGTAAGCAGTTCTGCACTTGCGTTCAATAT 1140
 Db 1081 GAAATGCTAAACAGGTTGATTCACATTTGTAAGCAGTTCTGCACTTGCGTTCAATAT 1140
 QY 1141 GGCATGATGCTGACGCTGCTCCAAAGATGCTCATTTGCTGATGGATATGGCTCCCA 1200
 Db 1141 GGCATGATGCTGACGCTGCTCCAAAGATGCTCATTTGCTGATGGATATGGCTCCCA 1200
 QY 1201 ACTGACATTTACCTCCGAGAGGCTGTGTGCGACACGAGCTGCTCATTTGATGAGC 1260
 Db 1201 ACTGACATTTACCTCCGAGAGGCTGTGTGCGACACGAGCTGCTCATTTGATGAGC 1260
 QY 1261 TTTGATGACAGATTCATTTGCTGATCTGGAGATGAGACTTAAAGSTATGGAACACA 1320
 Db 1261 TTTGATGACAGATTCATTTGCTGATCTGGAGATGAGACTTAAAGSTATGGAACACA 1320
 QY 1321 AGTACTTGTGAATTTGTAAAGACCTTAATGACACAAAGAGGCTGCTGTTTGGCAG 1380
 Db 1321 AGTACTTGTGAATTTGTAAAGACCTTAATGACACAAAGAGGCTGCTGTTTGGCAG 1380
 QY 1381 TACAGGAGACAGCTGTAGTGTGCTCATCTGACACATCATCATGATGATGAGACATA 1440
 Db 1381 TACAGGAGACAGCTGTAGTGTGCTCATCTGACACATCATCATGATGATGAGACATA 1440
 QY 1441 GAATGTGTGATGTTTAAAGGATGTAAGAGGATGAGAGATGGTGGTGTATGCA 1500
 Db 1441 GAATGTGTGATGTTTAAAGGATGTAAGAGGATGAGAGATGGTGGTGTATGCA 1500
 QY 1501 TTTGATTAACAGAGATGATGATGAGGCTTATGATGAAAAATTAATGTGGGATCTT 1560
 Db 1501 TTTGATTAACAGAGATGATGATGAGGCTTATGATGAAAAATTAATGTGGGATCTT 1560
 QY 1561 GTGGCTGCTTTTGGACCCCGTGTCTCTGCAAGGACACTGTGTCTACGACCTTGTGGAG 1620
 Db 1561 GTGGCTGCTTTTGGACCCCGTGTGTCTCTGCAAGGACACTGTGTCTACGACCTTGTGGAG 1620
 QY 1621 CATTCGGAAGAGTTTGTGACTACAGTTGATGAAATTCACAGTTTGTGCTCAGTTTCACAT 1680
 Db 1621 CATTCGGAAGAGTTTGTGACTACAGTTGATGAAATTCACAGTTTGTGCTCAGTTTCACAT 1680

QY 1681 GATGACACATTCATTCATGAGGACTTCCTAATATATCATCAGCTGCCAAGCTGAACCCCC 1740
 Db 1681 GATGACACATTCATTCATGAGGACTTCCTAATATATCATCAGCTGCCAAGCTGAACCCCC 1740
 QY 1741 CGTTCCTCTTCGAACTACACCTTACATTCGAGATTAATTAACCATACACTGACTCAT 1800
 Db 1741 CGTTCCTCTTCGAACTACACCTTACATTCGAGATTAATTAACCATACACTGACTCAT 1800
 QY 1801 ACTTCCCGAGACCATTAAGTTGGGATTTTAAAGTATGATGATGCAATACAGGATGAGC 1860
 Db 1801 ACTTCCCGAGACCATTAAGTTGGGATTTTAAAGTATGATGATGCAATACAGGATGAGC 1860
 QY 1861 AACACAGTAACTAACTACTGCTCCAGATTTCTCCAGATACCGAGAGAGGCTT 1920
 Db 1861 AACACAGTAACTAACTACTGCTCCAGATTTCTCCAGATACCGAGAGAGGCTT 1920
 QY 1921 TGAGACTCTGTTGGGACACAGTTGTCTGCACTGCGGCCAGAGAGGCTTACTACAGACA 1980
 Db 1921 TGAGACTCTGTTGGGACACAGTTGTGTCTGCACTGCGGCCAGAGAGGCTTACTACAGACA 1980
 QY 1981 ACTGACTCTTCAGTGTGCTATCAGAAAGATGCTCTATCATATTTGAAATGATGGAAC 2040
 Db 1981 ACTGACTCTTCAGTGTGCTATCAGAAAGATGCTCTATCATATTTGAAATGATGGAAC 2040
 QY 2041 TTTTAAACCTCCCT 2100
 Db 2041 TTTTAAACCTCCCT 2100
 QY 2101 CAGACAAAGGAGCTTAAATATATTTAGTGTGTTTCCGAAAAA 2151
 Db 2101 CAGACAAAGGAGCTTAAATATATTTAGTGTGTTTCCGAAAAA 2151

RESULT 3
 AA293350
 ID AA293350 standard; cdna; 2151 bp.
 XX
 AC AA293350;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE Sequence encoding F-box protein FBP-1.
 XX
 KW F-box protein; FBP; diagnosis; treatment; screening; agonist;
 KW antagonist; proliferative disorder; differential disorder;
 KW breast cancer; prostate cancer; ovarian cancer; cancer;
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
 KW inflammatory disorder; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 70..1779
 FT CDS /*tag= a
 FT /product= FBP-1
 PN W0200012679-A1.
 XX
 PD 09-MAR-2000.
 XX
 PE 27-AUG-1999; 99MO-US19560.
 XX
 PR 28-AUG-1998; 98US-0098355.
 PR 03-FEB-1999; 99US-0118568.
 PR 15-MAR-1999; 99US-0124449.
 XX
 PA (UNY) UNIV NEW YORK STATE.
 XX
 PI Chlaure DS, Pagano M, Latres E;
 XX
 DR WPI: 2000-256635/22.
 DR P-PSDB; AAT83041.

xx Novel nucleic acid for screening compounds useful for treating
 PT proliferative and differentiative disorders such as cancer and immune
 PT disorders comprises sequences encoding ubiquitin ligases
 xx
 PS Disclosure; Figure 3b, 3c, 3d; 245pp; English.
 CC
 CC Nucleic acids encoding substrate-targeting subunits of ubiquitin
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis
 CC of proliferative and differentiated related disorders by measuring
 CC FBP gene expression. Cells expressing such proteins or
 CC their fragments are useful for screening compounds. The compounds
 CC are agonists or antagonists, which are useful for treating a
 CC proliferative or differentiative disorder in a mammal such as
 CC breast, ovarian and prostate cancer and small cell lung carcinoma
 CC and also major opportunistic infections, immune disorders,
 CC cardiovascular diseases and inflammatory disorders. FBP protein,
 CC analogs, derivatives and their subsequences, anti-FBP antibodies
 CC are also useful in diagnosis of the disorders.
 CC
 SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;
 Query Match 100.0%; Score 2151; DB 21; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
 Matches 2151; Conservative 0; Indels 0; Gaps 0;
 QY 1 TCGTGTGGCTGGCGCTGGACCAAGGCGCGCCCGCGGAGAGCGGACCACTGGCC 60
 DB 1 TCGGTGTGGCTGGCGCTGGACCAAGGCGCGCCCGCGGAGAGCGGACCACTGGCC 60
 QY 61 TCGCGCATTTAGGAGCCCGCGCGCGGCTGTGCAAGAGAGAGCCCTCAAGTTTATGAT 120
 DB 61 TCGCGCATTTAGGAGCCCGCGCGCGGCTGTGCAAGAGAGAGCCCTCAAGTTTATGAT 120
 QY 121 TCCCTAG 180
 DB 121 TCCCTAG 180
 QY 181 AATTCAGTACAG 240
 DB 181 AATTCAGTACAG 240
 QY 241 TGTTTAG 300
 DB 241 TGTTTAG 300
 QY 301 GGCACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 DB 301 GGCACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 QY 361 AAGGAAGTGTGTCAAAATCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 DB 361 AAGGAAGTGTGTCAAAATCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 421 GAACATCTTATATCCCAATGTCATTAACCAATGTCATTAACCAATGTCATTAACCA 480
 DB 421 GAACATCTTATATCCCAATGTCATTAACCAATGTCATTAACCAATGTCATTAACCA 480
 QY 481 CCTATGTGAGAGAGATTTCAATGCTGTGCGAGCTCGGGGATGATCATATCGCT 540
 DB 481 CCTATGTGAGAGAGATTTCAATGCTGTGCGAGCTCGGGGATGATCATATCGCT 540
 QY 541 GAGAACATTTCTGTCAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 DB 541 GAGAACATTTCTGTCAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 601 GATGTACCGAG 660
 DB 601 GATGTACCGAG 660
 QY 661 AGGACAGATTTCTGTGAG 720
 DB 661 AGGACAGATTTCTGTGAG 720

QY 721 AAAACAACTCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACACTTTATCC 780
 DB 721 AAAACAACTCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACACTTTATCC 780
 QY 781 AAAATTTACAG 840
 DB 781 AAAATTTACAG 840
 QY 841 CAGAGATTTACAG 900
 DB 841 CAGAGATTTACAG 900
 QY 901 CAGAGATTTACAG 960
 DB 901 CAGAGATTTACAG 960
 QY 961 GAATGCAAGGAG 1020
 DB 961 GAATGCAAGGAG 1020
 QY 1021 AGAGTATCATTAACAGAGATCATGATTCACAGGTCAGAGTGTGGATGTAATACAGGT 1080
 DB 1021 AGAGTATCATTAACAGAGATCATGATTCACAGGTCAGAGTGTGGATGTAATACAGGT 1080
 QY 1081 GAAATGCTAAACAG 1140
 DB 1081 GAAATGCTAAACAG 1140
 QY 1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 DB 1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 QY 1201 ACTGACATTTACCTCCGAG 1260
 DB 1201 ACTGACATTTACCTCCGAG 1260
 QY 1261 TTTGATGACAG 1320
 DB 1261 TTTGATGACAG 1320
 QY 1321 AGTACTGTGATTTGTAAG 1380
 DB 1321 AGTACTGTGATTTGTAAG 1380
 QY 1381 TACAGGAG 1440
 DB 1381 TACAGGAG 1440
 QY 1441 GAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 DB 1441 GAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 QY 1501 TTTGATTAACAG 1560
 DB 1501 TTTGATTAACAG 1560
 QY 1561 GTGCTGCTTTGAG 1620
 DB 1561 GTGCTGCTTTGAG 1620
 QY 1621 CATTTCCGAG 1680
 DB 1621 CATTTCCGAG 1680
 QY 1681 GATGACAG 1740
 DB 1681 GATGACAG 1740
 QY 1741 GGTTCCTCTTCGAG 1800
 DB 1741 GGTTCCTCTTCGAG 1800


```

Db      ||| 840
781 AAAATATACAGACATTGAGACAAATAGAAATCTAATGGAGATGTGGAAGCAATAGTTTA
QY      ||| 900
841 CAGAAATTCATGCGGAGTGAAGCAAGCAAGAGATTTCTGTTTACAGTATGATGAT
Db      ||| 900
841 CAGAAATTCATGCGGAGTGAAGCAAGCAAGAGATTTCTGTTTACAGTATGATGAT
QY      ||| 960
901 CAGAAATATAGTAAGCGGCTTGCAGACAAACAATCAAGATCTGGATTAACACATTTG
Db      ||| 960
901 CAGAAATATAGTAAGCGGCTTGCAGACAAACAATCAAGATCTGGATTAACACATTTG
QY      ||| 1020
961 GAATGCAAGCAATTTCTACAGGCGCATACAGTTCAGTCCCTGCTCCAGTATGATGAG
Db      ||| 1020
961 GAATGCAAGCAATTTCTACAGGCGCATACAGTTCAGTCCCTGCTCCAGTATGATGAG
QY      ||| 1080
1021 AGAGTGCATACAGATCATCGGATTCACAGGTCAGAGTGGGATGTAATATACAGGT
Db      ||| 1080
1021 AGAGTGCATACAGATCATCGGATTCACAGGTCAGAGTGGGATGTAATATACAGGT
QY      ||| 1140
1081 GAAATGCTAAACACGTTGATTCACCATTTGTAGACAGTTTGCACCTTGCTTCATPAT
Db      ||| 1140
1081 GAAATGCTAAACACGTTGATTCACCATTTGTAGACAGTTTGCACCTTGCTTCATPAT
QY      ||| 1200
1141 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      ||| 1200
1141 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
QY      ||| 1260
1201 ACTGACATTTACGCTCGGAGGCTGCTGCGACACGAGTGTGTCATGTTGATGAC
Db      ||| 1260
1201 ACTGACATTTACGCTCGGAGGCTGCTGCGACACGAGTGTGTCATGTTGATGAC
QY      ||| 1320
1261 TTTGATGACAAAGTACATTTGTTGATGATGATGATGATGATGATGATGATGATGAT
Db      ||| 1320
1261 TTTGATGACAAAGTACATTTGTTGATGATGATGATGATGATGATGATGATGATGAT
QY      ||| 1380
1321 AGTACTGTTGAATTTGTAAAGACCTTAATGAGACCAACAGGATGCTGTTTGAG
Db      ||| 1380
1321 AGTACTGTTGAATTTGTAAAGACCTTAATGAGACCAACAGGATGCTGTTTGAG
QY      ||| 1440
1381 TACAGGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      ||| 1440
1381 TACAGGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
QY      ||| 1500
1441 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      ||| 1500
1441 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
QY      ||| 1560
1501 TTTGATTAACAGAGATAGTACAGTGGGCTATGATGAAAAATTAAGTGGGATCTT
Db      ||| 1560
1501 TTTGATTAACAGAGATAGTACAGTGGGCTATGATGAAAAATTAAGTGGGATCTT
QY      ||| 1620
1561 GTGGCTGTTTGAGACCCCGGCTGCTGACAGGACACTCTGTCTACGAGACCTTGTGAG
Db      ||| 1620
1561 GTGGCTGTTTGAGACCCCGGCTGCTGACAGGACACTCTGTCTACGAGACCTTGTGAG
QY      ||| 1680
1621 CATTCGGAAGATTTTTCAGACTACAGTTTGTAGTAATTCAGATTTGTACAT
Db      ||| 1680
1621 CATTCGGAAGATTTTTCAGACTACAGTTTGTAGTAATTCAGATTTGTACAT
QY      ||| 1740
1681 CATGACAAATCTATCTGGGACTTCTTAATGATCCAGGCTGCCAAGCTGAACCCCC
Db      ||| 1740
1681 CATGACAAATCTATCTGGGACTTCTTAATGATCCAGGCTGCCAAGCTGAACCCCC
QY      ||| 1800
1741 CTTTCCCTTCGAAATACACATCTACATCTCCAGATTAATACATACATGACCTCAT
Db      ||| 1800
1741 CTTTCCCTTCGAAATACACATCTACATCTCCAGATTAATACATACATGACCTCAT
QY      ||| 1860
1801 ACTTCCCAAGACCCATTAAGTGGGATTTAAAGTATCTGCAATATACAGATGAGC
Db      ||| 1860
1801 ACTTCCCAAGACCCATTAAGTGGGATTTAAAGTATCTGCAATATACAGATGAGC
QY      ||| 1920
1861 AACACAGTAAACAACTACTGCGGAGTTCCTGAGTACAGCCGAGACAGGCGTT

```

```

Db      ||| 1920
1861 AACACAGTAAACAACTACTGCGGAGTTCCTGAGTACAGCCGAGACAGGCGTT
QY      ||| 1980
1921 TGAGACTCTGTTGGGACACAGTGTGCTGACAGTCCGCGAGAGCGGTACTACACACA
Db      ||| 1980
1921 TGAGACTCTGTTGGGACACAGTGTGCTGACAGTCCGCGAGAGCGGTACTACACACA
QY      ||| 2040
1981 ACTGACTGCTTTCAGTGTGCTATCAGAAAGTGTCTTCTATCAATTTGATGATGGAAC
Db      ||| 2040
1981 ACTGACTGCTTTCAGTGTGCTATCAGAAAGTGTCTTCTATCAATTTGATGATGGAAC
QY      ||| 2100
2041 TTTTAAACCTCCCTCCTCCTCCTTTCACCTCTGACCTATTTTCCATTTGGTTC
Db      ||| 2100
2041 TTTTAAACCTCCCTCCTCCTCCTTTCACCTCTGACCTATTTTCCATTTGGTTC
QY      ||| 2151
2101 CAGACAAAGGTGACTTAAATATATTTAGTGTGTCGAGAAAAA
Db      ||| 2151
2101 CAGACAAAGGTGACTTAAATATATTTAGTGTGTCGAGAAAAA
RESULT 5
AAC84610
ID AAC84610 standard; DNA; 2151 BP.
AAC84610:
02-APR-2001 (first entry)
Human ZF1 protein encoding DNA.
S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
Bad; Bcl-2; tumour; cytosolic; ds.
Homo sapiens.
W0200075184-A1.
14-DEC-2000.
05-JUN-2000; 2000MO-US15449.
04-JUN-1999; 990S-0137494.
(UYUA) UNIV YALE.
Zhang H, Tsvetkov LM, Kondo T;
WPI; 2001-061703/07.
P-PSDB; AAB48298.
Modulating polypeptide levels in a cell, diagnosing and treating tumor,
PT involves altering levels of proteins such as S-phase kinase associated
PT proteins 1, 2 and cullin/CDC53 proteins -
XX
PS Examples: Page 129-130; 162pp; English.
XX
CC The invention relates to methods of altering the polypeptide levels in a
CC cell, using proteins selected from S-phase kinase associated proteins 1
CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
CC cullin/CDC53 family of proteins). The method is useful for altering the
CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents
CC that modulate interactions between SKP and target proteins are useful for
CC treating tumours.
SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;
Query Match 100.0%; Score 2151; DB 22; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCCTTGCTGCGGCTGCGACCAAGAGGCGGCCCGCGGAGAGCGGACCGACGATGGCC 60

```


ID AH90079 standard; CDNA; 3220 BP.
AC AAH90079;
XX 01-OCT-2001 (first entry)
DE Human bone marrow CDNA, SEQ ID NO: 323.
XX
XX Human; bone marrow; antinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX Homo sapiens.
OS
XX MO200153453-A2.
XX
XX 26-JUL-2001.
PD
XX 23-DEC-2000; 2000WO-US34960.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0634350.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
P1 Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
P1 Zhao QA, Zhou P, Drmanac RT;
XX WPI: 2001-488707/53.
DR P-PSDB: AAM00960.
XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 428; 648pp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
CC expressed in the bone marrow. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
XX Sequence 3220 BP; 878 A; 725 C; 770 G; 847 T; 0 other;
SQ
Query Match 98.6%; Score 2120.8; DB 22; Length 3220;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
OY 1 TGGCTTGGCTGGCGGCAACAAAGGGGGGGCCCGGGGAGAGGCGACCGAGTGGCC 60
DB 59 TGGCTTGGCTGGCGGCTGGCAACAAAGGGGGGGCCCGGGGAGAGGCGACCGAGTGGCC 118
OY 61 TGGGCGATTATGAGACCGCGGCGAGGGGCTGCTGCAAGAGAGGACCTCAAGTTATGAT 120
DB 119 TGGGCGATTATGAGACCGCGGCGAGGGGCTGCTGCAAGAGAGGACCTCAAGTTATGAT 178

OY 121 TCCTCAGAGAGAGAGACTGTAAATTAATGGGAACCCCTAGAGAGATTAATACCAGAGAG 180
DB 179 TCCTCAGAGAGAGAGAGACTGTAAATTAATGGGAACCCCTAGAGAGATTAATACCAGAGAG 238
OY 181 AATTTCAGTACAGACATTCACACAGCTGTCCAGACTGTGCTTAACCAAGAACAGTA 240
DB 239 AATTTCAGTACAGACATTCACACAGCTGTCCAGACTGTGCTTAACCAAGAACAGTA 298
OY 241 TGTTCAGACAGCACTGCTATGAGAGACTGAGAAATTTGTGGCCAAACAAACACTGGCAAT 300
DB 299 TGTTCAGACAGCACTGCTATGAGAGACTGAGAAATTTGTGGCCAAACAAACACTGGCAAT 358
OY 301 GGCAGCTTCAGATATGATTTGTGCCAAGCAAGAAACTCTCAGACAGCTATGAGAAAGAA 360
DB 359 GGCAGCTTCAGATATGATTTGTGCCAAGCAAGAAACTCTCAGACAGCTATGAGAAAGAA 418
OY 361 AAGGAAGTGTGTGCAAAATCTTTGAGAGAGTGTGCTGAGAGTCAAGTCAAGTATGAG 420
DB 419 AAGGAAGTGTGTGCAAAATCTTTGAGAGAGTGTGCTGAGAGTCAAGTCAAGTATGAG 478
OY 421 GAACATCTATATCCCAATATGCTCATTAACCAACATGAGGACATTAACCTGTATCTTAA 480
DB 479 GAACATCTATATCCCAATATGCTCATTAACCAACATGAGGACATTAACCTGTATCTTAA 538
OY 481 CTTATGTTGCAAGAGATTTTCAATTAAGTCTGCTGCAAGCTGGGAGTTGATCATATGCT 540
DB 539 CTTATGTTGCAAGAGATTTTCAATTAAGTCTGCTGCAAGCTGGGAGTTGATCATATGCT 598
OY 541 GAGAACATCTGTCATATCCAGTGGCAATCAATCAATGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 599 GAGAACATCTGTCATATCCAGTGGCAATCAATCAATGCTGCTGCTGCTGCTGCTGCTGCT 658
OY 601 GAATGTACAGAGTGAAGCTGTGATGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTG 660
DB 659 GAATGTACAGAGTGAAGCTGTGATGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTG 718
OY 661 AGGACAGATTTCTGTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTG 720
DB 719 AGGACAGATTTCTGTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTG 778
OY 721 AAAAAGCAACCTCTGTGAGAGAGGAGTGTGCAACCTCTTTTATGAGAGAGTGTGCT 780
DB 779 AAAAAGCAACCTCTGTGAGAGAGGAGTGTGCAACCTCTTTTATGAGAGAGTGTGCT 838
OY 781 AAAATTATACAGACATTTGAGACATTAATGATTAATGAGATGTGAGACATTAATTTA 840
DB 839 AAAATTATACAGACATTTGAGACATTAATGATTAATGAGATGTGAGACATTAATTTA 898
OY 841 CAGAGAAATTCAGTGTGCGAAGTGAACCAAGCAAGAGAGTGTGAGAGTGTGAT 900
DB 899 CAGAGAAATTCAGTGTGCGAAGTGAACCAAGCAAGAGAGTGTGAGAGTGTGAT 958
OY 901 CAGAGAAATTCAGTGTGCGAAGTGAACCAAGCAAGAGTGTGAGAGTGTGAT 960
DB 959 CAGAGAAATTCAGTGTGCGAAGTGAACCAAGCAAGAGTGTGAGAGTGTGAT 1018
OY 961 GAATGCAAGGAAATTTCTACAGAGCCATACAGTTCAGTGTGCTGCTGAGTATGATGAG 1020
DB 1019 GAATGCAAGGAAATTTCTACAGAGCCATACAGTTCAGTGTGCTGCTGAGTATGATGAG 1078
OY 1021 AGAGTATTCATTAACAGATATGAGATTCAGGCTCAGAGTGTGAGATTAATACAGGT 1080
DB 1079 AGAGTATTCATTAACAGATATGAGATTCAGGCTCAGAGTGTGAGATTAATACAGGT 1138
OY 1081 GAATGCTTAACACGTTGATTCACATGTTGAAGAGCTGTGCACTTGGCTTCAATTAAT 1140
DB 1139 GAATGCTTAACACGTTGATTCACATGTTGAAGAGCTGTGCACTTGGCTTCAATTAAT 1198
OY 1141 GGCATGATGTGAGACTGCTCAAAAGATCGTTCCATTGCTGTATGAGGATATGGCTCCCA 1200
DB 1199 GGCATGATGTGAGACTGCTCAAAAGATCGTTCCATTGCTGTATGAGGATATGGCTCCCA 1258

QY 1201 ACTGACATTACCTCCGGAGGCTGCTGTCGACACGAGCTCTGTCATGTTGTAGAC 1260
 DB 1259 ACTGACATTACCTCCGGAGGCTGCTGTCGACACGAGCTCTGTCATGTTGTAGAC 1318
 QY 1261 TTTGATGACAAAGTACATTTGTTTCTGCATCTGGGGATAGAACTTAAAGCTATGGAACAA 1320
 DB 1319 TTTGATGACAAAGTACATTTGTTTCTGCATCTGGGGATAGAACTTAAAGCTATGGAACAA 1378
 QY 1321 AGTACTTGTCAATTTGTTAAGGACCTTAAATGACACAAAGGAGCTGCTGTTTGCAG 1380
 DB 1379 AGTACTTGTCAATTTGTTAAGGACCTTAAATGACACAAAGGAGCTGCTGTTTGCAG 1438
 QY 1381 TACAGGACAGGCTGTGAGTGTGCTCATCTGACAAACATATGATTATGAGCAATA 1440
 DB 1439 TACAGGACAGGCTGTGAGTGTGCTCATCTGACAAACATATGATTATGAGCAATA 1498
 QY 1441 GAAATGTGTGATGTTTACAGTGTGAGAAAGGACATGAGGAATGTGCTGTTATTCGA 1500
 DB 1499 GAAATGTGTGATGTTTACAGTGTGAGAAAGGACATGAGGAATGTGCTGTTATTCGA 1558
 QY 1501 TTTGATTAACAAGAGATAGTACATGGGGCTATGATGAAAAATTAAGTGTGGATCTT 1560
 DB 1559 TTTGATTAACAAGAGATAGTACATGGGGCTATGATGAAAAATTAAGTGTGGATCTT 1618
 QY 1561 GTGGCTGCTTTGAGACCCGCTGCTCGAGAGGACACTGTGCTACGAGACCTTGTGAG 1620
 DB 1619 GTGGCTGCTTTGAGACCCGCTGCTCGAGAGGACACTGTGCTACGAGACCTTGTGAG 1678
 QY 1621 CATTTCCGGAAGATTTTTCGACTACAGTTTGATGAATTCGACATGTCAGTATTCACAT 1680
 DB 1679 CATTTCCGGAAGATTTTTCGACTACAGTTTGATGAATTCGACATGTCAGTATTCACAT 1738
 QY 1681 GATGACACATCTCTATGTCGGGACTCTAAATGATTCAGGCTGAGCCCACTGAACCCGCC 1740
 DB 1739 GATGACACATCTCTATGTCGGGACTCTAAATGATTCAGGCTGAGCCCACTGAACCCGCC 1798
 QY 1741 CGTTCCCTCTTGGAAATACACCTACATCTCCAGATAAATAACATACATGACCTCAT 1800
 DB 1799 CGTTCCCTCTTGGAAATACACCTACATCTCCAGATAAATAACATACATGACCTCAT 1858
 QY 1801 ACTTGGCCAGGAGCCCAATTAAG--TTGCGGTATTTAAGCTATGTCGCCAATACCAAGATGAG 1859
 DB 1859 ACTTGGCCAGGAGCCCAATTAAGTTTGGGATTTAAGTATGTCGCCAATACCAAGATGAG 1918
 QY 1860 CAACAACAGTAAACATCAACTACTGCCAGTTTCCCTGGAGTAGCCGAGAGCAAGGCT 1919
 DB 1919 CAACAACAGTAAACATCAACTACTGCCAGTTTCCCTGGAGTAGCCGAGAGCAAGGCT 1978
 QY 1920 TTTAGACTCCTGTGGGACACAGTGTGTCTGACAGTGGCCCGAGAGGCTCTACTACAGAC 1979
 DB 1979 TTTAGACTCCTGTGGGACACAGTGTGTCTGACAGTGGCCCGAGAGGCTCTACTACAGAC 2038
 QY 1980 AACTGACTCCTTACAGTGTGTATCAGAAAGATGTCTTATCAATTTGTAAGATTTGGA 2039
 DB 2039 AACTGACTCCTTACAGTGTGTATCAGAAAGATGTCTTATCAATTTGTAAGATTTGGA 2098
 QY 2040 CTTTAAACCTCCCTCCTCCTCCTTCCACTGACACCTAGTATTTTCCATTTGTT 2099
 DB 2099 CTTTAAACCTCCCTCCTCCTCCTTCCACTGACACCTAGTATTTTCCATTTGTT 2158
 QY 2100 CCAGCAAAAGTGACTTATTAATATTTAGTGTGTTTGCACAA 2143
 DB 2159 CCAGCAAAAGTGACTTATTAATATTTAGTGTGTTTGCACAA 2202

DE Human cell signalling protein-12 encoding cDNA.
 XX
 KW Cell signalling protein-12; CSIGP-12; cell proliferation;
 KW inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;
 KW arteriosclerosis; Addison's disease; multiple sclerosis; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 70..1779
 FT /tag= a
 FT /product= "Cell Signalling protein-12"
 XX
 PN W0958558-A2.
 XX
 PD 18-NOV-1999.
 XX
 PF 13-MAY-1999; 99WO-US10567.
 XX
 PR 13-MAY-1998; 98US-0085343.
 PR 26-AUG-1998; 98US-0098010.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;
 PI Baughn MR, Yang J;
 XX
 DR WPI: 2000-086432/07.
 DR P-PSDB: AAY44249.
 XX
 PT Human cell signalling proteins useful for, e.g. diagnosing cell
 PT proliferative and inflammatory disorders
 PS
 PS Claim 9; Page 87-88; 90pp; English.
 XX
 CC The present sequence is a cDNA obtained from Incyte clone 3239149 of
 CC COVADT01 library. It encodes cell signalling protein-12 (CSIGP-12). It
 CC is expressed in musculo-skeletal, gastrointestinal and nervous
 CC tissues. Fragments of CSIGP encoding nucleic acid can be used as
 CC hybridisation probe for detecting CSIGP related sequences or allelic
 CC variants. Recombinant CSIGP can be produced in host cells by transforming
 CC them with genetically engineered vectors. Agonists or antagonists can be
 CC used in the treatment of cell proliferative and inflammatory disorders
 CC associated with decreased or increased CSIGP expression. CSIGP is used in
 CC the diagnosis, prevention and treatment of cell proliferative disorders
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.
 XX
 SO Sequence 2419 BP; 671 A; 531 C; 625 G; 590 T; 2 other;
 Query Match 98.3%; Score 2114.8; DB 21; Length 2419;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2119; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 11 GGGGCTTGACCAAAAGGGGCGCCCGGAGAGGAGCCAGTGGCTCGGCGATTA 70
 DB 11 GGGGCTTGACCAAAAGGGGCGCCCGGAGAGGAGCCAGTGGCTCGGCGATTA 70
 QY 71 TGGACCCGGCCGAGGCGGTGCTGCAAGAAAGGACATCAAGTTATGAATTCCTCAGAGA 130
 DB 71 TGGACCCGGCCGAGGCGGTGCTGCAAGAAAGGACATCAAGTTATGAATTCCTCAGAGA 130
 QY 131 GAGAACTGTAAATATGCGAACCCTTGAAGATATACAGAGAAATTCATTA 190
 DB 131 GAGAACTGTAAATATGCGAACCCTTGAAGATATATACAGAGAAATTCATTA 190
 QY 191 GACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAAGATGTTTACAA 250
 DB 191 GACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAAGATGTTTACAA 250
 QY 251 GCACGTATGAAAGACATGAGAAATTTGTGGCCAAACAAACTTGCATATGGACATTCCA 310
 DB 251 GCACGTATGAAAGACATGAGAAATTTGTGGCCAAACAAACTTGCATATGGACATTCCA 310

RESULT 7
 AA229233
 ID AA229233 standard; cDNA; 2419 BP.
 XX
 AC AA229233;
 XX
 DT 28-FEB-2000 (first entry)
 XX

QY	311	GTATATATTGGCCCAAGACAGGAAACCTCAGCAACCTATGAAAAGGAAAGCAACCTCT	370
Db	311	GTATATATTGGCCCAAGACAGGAAACCTCTCAGCAACCTATGAAAAGGAAAGCAACCTCT	370
QY	371	GTGTCAAAATACTTTTGAGCAGTGGTCAAGATCAGATCAAGTGAATTTGGTGAACATCTTA	430
Db	371	GTGTCAAAATACTTTTGAGCAGTGGTCAAGATCAGATCAAGTGAATTTGGTGAACATCTTA	430
QY	431	TATGCCAAATGTGTCAATACCAATATGGGCACATAAACTGTATCTTAAACCTATGTTC	490
Db	431	TATGCCAAATGTGTCAATACCAATATGGGCACATAAACTGTATCTTAAACCTATGTGTTC	490
QY	491	AGAGAGATTTCATAACTGCTCTGACACTGGGGATTTGATTCATTCGGTGAACATTC	550
Db	491	AGAGAGATTTCATAACTGTCTCTGACACTGGGGATTTGATTCATTCGGTGAACATTC	550
QY	551	TGTCAATACCTGGATGCCAAATCATATGTCTGTGAACCTTTGTGTCAAGAAATGTATCC	610
Db	551	TGTCAATACCTGGATGCCAAATCATATGTCTGTGAACCTTTGTGTCAAGAAATGTATCC	610
QY	611	GAGTACACCTGTATGGCATGTCTGTGAAGAAGCTTATCGAGACAAATGTCTCAGACAGATT	670
Db	611	GAGTACACCTGTATGGCATGTCTGTGAAGAAGCTTATCGAGACAAATGTCTCAGACAGATT	670
QY	671	CTCTGTGGAGAGCCCTGGCAGAACGAAGAGATGGGCAGATTTTATTCAAAACCAAC	730
Db	671	CTCTGTGGAGAGCCCTGGCAGAACGAAGAGATGGGCAGATTTTATTCAAAACCAAC	730
QY	731	CTCCGAGCGGAATGTCTCTCCCAACCTTTTATAGACACTTATCCTAAATATTATAC	790
Db	731	CTCCGAGCGGAATGTCTCTCCCAACCTTTTATAGACACTTATCCTAAATATTATAC	790
QY	791	AAGACATTGAGACATAGAAATCTAATTGGAGATGTGGAACATACTTATACAGACAATTC	850
Db	791	AAGACATTGAGACATAGAAATCTAATTGGAGATGTGGAACATACTTATACAGACAATTC	850
QY	851	ACTGCCAAGTGAACCAAGCAAGAGATTTACTGTTTACAGTATGATATCGAATAATAG	910
Db	851	ACTGCCAAGTGAACCAAGCAAGAGATTTACTGTTTACAGTATGATATCGAATAATAG	910
QY	911	TAAAGGCGCTTCGAGACACACATCAATCAAGTCTGGATTAATAACACATTTGGAATCGAACG	970
Db	911	TAAAGGCGCTTCGAGACACACATCAATCAAGTCTGGATTAATAACACATTTGGAATCGAACG	970
QY	971	GAAATTCACAGGCCATACAGGTTCAGTCTCTGTCTCCAGATATGATGAGAGATGATCA	1030
Db	971	GAAATTCACAGGCCATACAGGTTCAGTCTCTGTCTCCAGATATGATGAGAGATGATCA	1030
QY	1031	TAAACAGATCTATCGGATTCACAGGTCAAGATGTGGGATTAATACAGGTGAATGCTTA	1090
Db	1031	TAAACAGATCTATCGGATTCACAGGTCAAGATGTGGGATTAATACAGGTGAATGCTTA	1090
QY	1091	ACACGTTGATTCACACATTTGTGAACAGATTCGTCACTTGGTTCAATATGACATGATGG	1150
Db	1091	ACACGTTGATTCACACATTTGTGAACAGATTCGTCACTTGGTTCAATATGACATGATGG	1150
QY	1151	TGACCTGTCTCCAAAGATTCCTTCATTCGTGTATGGGATATGGCTCCCACTGCATTTA	1210
Db	1151	TGACCTGTCTCCAAAGATTCCTTCATTCGTGTATGGGATATGGCTCCCACTGCATTTA	1210
QY	1211	CCCTCCGGAGGGTGTCTGTGGACACCGAGCTCTCTCAATTTGTAGACTTTGATGACA	1270
Db	1211	CCCTCCGGAGGGTGTCTGTGGACACCGAGCTCTCTCAATTTGTAGACTTTGATGACA	1270
QY	1271	AGTACATTTGTTTCGCATCTGGGGATAGAACTTAAGGTATGAGAACACAGTACTTGTG	1330
Db	1271	AGTACATTTGTTTCGCATCTGGGGATAGAACTTAAGGTATGAGAACACAGTACTTGTG	1330
QY	1331	AATTTGTAAAGACCTTAAATAGACACAAACGAGCATTTCTGTTTGCAGTACAGGGACA	1390
Db	1331	AATTTGTAAAGACCTTAAATAGACACAAACGAGCATTTCTGTTTGCAGTACAGGGACA	1390

OY	1391	GGCTGAGTAGTAGGCGCTCATCTTCGACAAACATCATCGATTTATGGACATATGAATGGGCG	1450		
OY	1391	GGCTGAGTAGTAGGCGCTCATCTTCGACAAACATCATCGATTTATGGACATATGAATGGGCG	1450		
Db	1391	GGCTGAGTAGTAGGCGCTCATCTTCGACAAACATCATCGATTTATGGACATATGAATGGGCG	1450		
OY	1451	CATGTTTACAGAGTGTATAGAAAGGCCATGAGAAATTTGGTGCCTTGTATTTCGATTGGATTAACA	1510		
Db	1451	CATGTTTACAGAGTGTATAGAAAGGCCATGAGAAATTTGGTGCCTTGTATTTCGATTGGATTAACA	1510		
OY	1511	AGAGGATGATGACAGTGGGGCCCTATGATGGAATAAATTAAAGTGTGGGATCTTGTGGCTGCTT	1570		
Db	1511	AGAGGATGATGACAGTGGGGCCCTATGATGGAATAAATTAAAGTGTGGGATCTTGTGGCTGCTT	1570		
OY	1571	TGAGACCCCGCGTCTCTGCGAGGACACTGTGCTACAGGACCCCTGTGGACATTTCCGGAA	1630		
Db	1571	TGAGACCCCGCGTCTCTGCGAGGACACTGTGCTACAGGACCCCTGTGGACATTTCCGGAA	1630		
OY	1631	GAGTTTTCGACATACAGTTTGAATGAATTCACAGATTTGTCAGTAGTTACATGATGACACAA	1690		
Db	1631	GAGTTTTCGACATACAGTTTGAATGAATTCACAGATTTGTCAGTAGTTACATGATGACACAA	1690		
OY	1691	TCTCATCTGGGACCTTCCTAAATGATGACGCTGCGCCAGCTGGAACCCCGCGTTCCCGCTT	1750		
Db	1691	TCTCATCTGGGACCTTCCTAAATGATGACGCTGCGCCAGCTGGAACCCCGCGTTCCCGCTT	1750		
OY	1751	CTCGACATACACCTACATCTCCAGATTAATTAACCATACATGACCTCATACTTGGCCAG	1810		
Db	1751	CTCGACATACACCTACATCTCCAGATTAATTAACCATACATGACCTCATACTTGGCCAG	1810		
OY	1811	GACCATTTAAAGTTGGGGGATTTAAACGTACTGCGCCAAACACAGATGACGACAAACACAGTA	1870		
Db	1811	GACCATTTAAAGTTGGGGGATTTAAACGTACTGCGCCAAACACAGATGACGACAAACACAGTA	1870		
OY	1871	ACAATCAAACTACTGGCCAGTTTCCCTGGACAGCCGAGAGCAGGGCTTTCGAGACTCCT	1930		
Db	1871	ACAATCAAACTACTGGCCAGTTTCCCTGGACAGCCGAGAGCAGGGCTTTCGAGACTCCT	1930		
OY	1931	GTTGGGACACAGTTGGTCTGCGAGTGGGCCGAGGAGGCTTACTCAGCACAACGACTGCT	1990		
Db	1931	GTTGGGACACAGTTGGTCTGCGAGTGGGCCGAGGAGGCTTACTCAGCACAACGACTGCT	1990		
OY	1991	TGACTGCTGATACAAAGATGCTCTTATCAATTTGTCGTAATGTAATGGAACCTTTTAAACCT	2050		
Db	1991	TGACTGCTGATACAAAGATGCTCTTATCAATTTGTCGTAATGTAATGGAACCTTTTAAACCT	2050		
OY	2051	CCCTCTCTCTCTCTCTTCACCTCGACACTAGTTTTTCCCATTTGGTTCACAGCAAAAG	2110		
Db	2051	CCCTCTCTCTCTCTCTTCACCTCGACACTAGTTTTTCCCATTTGGTTCACAGCAAAAG	2110		
OY	2111	TGACTTAAATATATTAGTGTATT 2136			
Db	2111	TGACTTAAATATATTAGTGTATT 2136			
RESULT 8					
ID	AAH89966	standard: cDNA; 3622 BP.			
AC	AAH89966;				
XX	01-OCT-2001	(first entry)			
XX	Human bone marrow cDNA, SEQ ID NO: 97.				
DE					
XX	Human; bone marrow; antiinflammatory; cytosolic; neuroprotective;				
KW	antiviral; antibacterial; antifungal; anti-HIV; haemostatic;				
KW	immunosuppressive; gene therapy; cytokine cell proliferation;				
KW	cell differentiation modulator; immune disorder; infection; cancer;				
KW	human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss				
OS	Homo sapiens.				
XX					
XX	W0200153453-A2.				

PD 26-JUL-2001.
 XX 23-DEC-2000; 2000MO-US34960.
 PF 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX (HYSE-) HYSEQ INC.
 PA Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 XX WPI: 2001-488707/53.
 DR P-PSDB; AAM00847.
 XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -
 PT Claim 1; Page 274-275; 648pp; English.
 PS The present sequence is one of 251 novel human polynucleotides
 CC expressed in the bone marrow. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC of increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 XX
 SQ Sequence 3622 BP; 961 A; 839 C; 874 G; 948 T; 0 other;
 Query Match 93.5%; Score 2011.4; DB 22; Length 3622;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2026; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Db 621 TATCTTAACCTATGTTGACAGAGATTTTCAATACCTCTGCCAGCTGGGGATTGGAT 680
 QY 532 CATATCGGTGAGAACTTCTCTATACCTGAGTCCAAATCACTATGCTGTGAACTT 591
 Db 681 CATATCGGTGAGAACTTCTCTATACCTGAGTCCAAATCACTATGCTGTGAACTT 740
 QY 592 GTGTGCAAGGAATGTTACCGAGTACCTCTGATGGGATGCTGTGGGAAGAGCTTATCGAG 651
 Db 741 GTGTGCAAGGAATGTTACCGAGTACCTCTGATGGGATGCTGTGGGAAGAGCTTATCGAG 800
 QY 652 AGAATGTCAGACAGATTCCTCTGTGAGAGGCTGTGAGAGAGAGAGATGGGAGAG 711
 Db 801 AGAATGTCAGACAGATTCCTCTGTGAGAGGCTGTGAGAGAGAGAGATGGGAGAG 860
 QY 712 TATTTATTTCAAAAACCAACCTCTGTGAGAGGAGTCTCTCCCACTCTTTTATAGCA 771
 Db 861 TATTTATTTCAAAAACCAACCTCTGTGAGAGGAGTCTCTCCCACTCTTTTATAGCA 920
 QY 772 CTTATCTCTAAATTTATACAAAGATTTGAGACATTAATCTAATGAGATGGAGAGA 831
 Db 921 CTTATCTCTAAATTTATACAAAGATTTGAGACATTAATCTAATGAGATGGAGAGA 980
 QY 832 CATAGTTTACAGAAATTTCACTGCGGAGTGAAGAAAGCAAGAGATTTACTGTTACAG 891
 Db 981 CATAGTTTACAGAAATTTCACTGCGGAGTGAAGAAAGCAAGAGATTTACTGTTACAG 1040
 QY 892 TATGATGATCAAAAATTTAGTAAGGCGCTTCGAGACACACATCAAGATCTGGATATA 951
 Db 1041 TATGATGATCAAAAATTTAGTAAGGCGCTTCGAGACACACATCAAGATCTGGATATA 1100
 QY 952 AACCATGGAATGCAAGCAAGTCTCAAGGCGATCAAGTTCAGTCTGTCTGACAG 1011
 Db 1101 AACCATGGAATGCAAGCAAGTCTCAAGGCGATCAAGTTCAGTCTGTCTGACAG 1160
 QY 1012 TATGATGAGAGATGATCATTAACAGATCATGAGATTCACAGTGTGGAGATGA 1071
 Db 1161 TATGATGAGAGATGATCATTAACAGATCATGAGATTCACAGTGTGGAGATGA 1220
 QY 1072 AATACAGGTGAATGCTTAAACAGCTGATGATTCACATGATGAGAGATTCGACCTGCT 1131
 Db 1221 AATACAGGTGAATGCTTAAACAGCTGATGATTCACATGATGAGAGATTCGACCTGCT 1280
 QY 1132 TTCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1191
 Db 1281 TTCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1340
 QY 1192 GCTTCCCACTGACATTAACCTCCGAGAGGCTGTGATGAGAGAGAGAGAGAGAGAG 1251
 Db 1341 GCTTCCCACTGACATTAACCTCCGAGAGGCTGTGATGAGAGAGAGAGAGAGAGAG 1400
 QY 1252 GTTGTAGACTTGTGATGACAACTGATGATGATGATGATGATGATGATGATGATGAT 1311
 Db 1401 GTTGTAGACTTGTGATGACAACTGATGATGATGATGATGATGATGATGATGATGAT 1460
 QY 1312 TGGAAACAGAGTACTGTTGAATTTGTGAAGACCTTAATGAGACAAACGAGCATTCGC 1371
 Db 1461 TGGAAACAGAGTACTGTTGAATTTGTGAAGACCTTAATGAGACAAACGAGCATTCGC 1520
 QY 1372 TGGTTCAGTACAGGATTA 1431
 Db 1521 TGGTTCAGTACAGGATTA 1580
 QY 1432 TGGGACATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1491
 Db 1581 TGGGACATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1640
 QY 1492 TGTATTCATTTGATTAACAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1551
 Db 1641 TGTATTCATTTGATTAACAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1700
 QY 1552 TGGGATCTTGGGAGGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1611

Db 1701 TGGGATCTTGGGCTGGCTTTGGAACCCCGCTGCTGCAAGGACACTCTGTCTACGGACC 1760
 QY 1612 CTGTGGAGCATCTCCGAGAGAGTTTTCAGTACAGTTTGATGATTCAGATTTGTCAGT 1671
 Db 1761 CTGTGGAGCATCTCCGAGAGAGTTTTCAGTACAGTTTGATGATTCAGATTTGTCAGT 1820
 QY 1672 AGTTACATGATGATGACAAATCCCAATCTGGAGTCTCTAAATATCCAGCTGCCAAGCT 1731
 Db 1821 AGTTACATGATGATGACAAATCCCAATCTGGAGTCTCTAAATATCCAGCTGCCAAGCT 1880
 QY 1732 GAAACCCCGCTGCTGCTGCAACATACACCTACATCTCCAGATTAATTAACATACAC 1791
 Db 1881 GAAACCCCGCTGCTGCTGCAACATACACCTACATCTCCAGATTAATTAACATACAC 1940
 QY 1792 TGACCTCATACCTGCTCCAGAGACCCATTAAG-TTGGGATATTAATTAACATGTCGCAATAC 1850
 Db 1941 TGACCTCATACCTGCTCCAGAGACCCATTAAGTTGGGATATTAATTAACATGTCGCAATAC 2000
 QY 1851 CAGATGAGCAACACACATGATACAAATCACTGCTCCAGATTTCCCTGACATACGCGAGG 1910
 Db 2001 CAGATGAGCAACACACATGATACAAATCACTGCTCCAGATTTCCCTGACATACGCGAGG 2060
 QY 1911 AGCAGGCTTTGAGACTCTGTTGGACACAGTTGCTGCAAGTGGCCGACAGAGTCT 1970
 Db 2061 AGCAGGCTTTGAGACTCTGTTGGACACAGTTGCTGCAAGTGGCCGACAGAGTCT 2120
 QY 1971 ACTGACCAACATGACTGCTTACAGTCTCTATACAGAGATGCTTATACATTTGTGA 2030
 Db 2121 ACTGACCAACATGACTGCTTACAGTCTCTATACAGAGATGCTTATACATTTGTGA 2180
 QY 2031 TGATTTGGAATTTTAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2090
 Db 2181 TGATTTGGAATTTTAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2240
 QY 2091 CCATTGGTTCCAGACAAAGTGACTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2143
 Db 2241 CCATTGGTTCCAGACAAAGTGACTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2293

RESULT 9
 AAK52699
 ID AAK52699 standard; cDNA; 3003 BP.
 XX
 AC AAK52699;
 XX
 DI 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 2228.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 19-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0683361.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Tang YF, Liu C, Dirmann RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 P-PSDB: AAM79556.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 XX useful in diagnosis and gene therapy -
 PS Claim 1: Page 4584-4585; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 SQ Sequence 3003 BP: 807 A; 703 C; 721 G; 772 T; 0 other;

Query Match 93.4%; Score 2008.2; DB 22; Length 3003;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 112 TTATGAAATTCCTCAGAGAGAGAAAGACTGTAATATGCGCAACCCCTAGAGATATA 171
 Db 298 TTCCAGAAATTCCTCAGAGAGAGAAAGACTGTAATATGCGCAACCCCTAGAGATATA 357
 QY 172 CCAGAGAAAGAAATTCCTCAGAGAGAGAAAGACTGTAATATGCGCAACCCCTAGAGATATA 231
 Db 358 CCAGAGAAAGAAATTCCTCAGAGAGAGAAAGACTGTAATATGCGCAACCCCTAGAGATATA 417
 QY 232 GAAACATGATGTTTACAGACAGCTGCTATGAGACAGTGAATTTGTGGCCAAACAAA 291
 Db 418 GAAACATGATGTTTACAGACAGCTGCTATGAGACAGTGAATTTGTGGCCAAACAAA 477
 QY 292 CTTGCAATGGCACTTCAGATGATGTTGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 351
 Db 478 CTTGCAATGGCACTTCAGATGATGTTGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 537
 QY 352 GAAAGGAAAGGAAAGCAAGCTGTGCAAAATTAATTTAGCAGTGTGCAAGTCAAGTGTG 411
 Db 538 GAAAGGAAAGGAAAGCAAGCTGTGCAAAATTAATTTAGCAGTGTGCAAGTCAAGTGTG 597
 QY 412 GAAATTTGGAACATCTTATATCCCAATGTGTCATTTACCAACATGGGCACATAACTCG 471
 Db 598 GAAATTTGGAACATCTTATATCCCAATGTGTCATTTACCAACATGGGCACATAACTCG 657
 QY 472 TATCTTAACCTATGTTGAGAGAGATTTATCAATGCTGCTGCGAGTGGGATGGAT 531
 Db 658 TATCTTAACCTATGTTGAGAGAGATTTATCAATGCTGCTGCGAGTGGGATGGAT 717
 QY 532 CATATGCTGAGAAACATTTCTGTCATACCTGAGATGCCAAATCACTATGCTGTAACCT 591
 Db 718 CATATGCTGAGAAACATTTCTGTCATACCTGAGATGCCAAATCACTATGCTGTAACCT 777
 QY 592 GTGTGCAAGGAATGTACAGAGTACCTGAGATGGATGCTGTGGAAGAAAGCTTATCGAG 651
 Db 778 GTGTGCAAGGAATGTACAGAGTACCTGAGATGGATGCTGTGGAAGAAAGCTTATCGAG 837
 QY 652 AGAATGTCAGACAGATTTCTGTGAGAGAGGCTGTGCAAGAAAGAGATGGGAGACAG 711
 Db 838 AGAATGTCAGACAGATTTCTGTGAGAGAGGCTGTGCAAGAAAGAGATGGGAGACAG 897

QY 712 TATTATTCAAAAACAACTCTGACGGAGATCTCTCCCAACTCTTTTATAGACA 771
 |||||||
 Db 712 TATTATTCAAAAACAACTCTGACGGAGATCTCTCCCAACTCTTTTATAGACA 957
 |||||||
 QY 772 CTTATCTCTAAATATATCAAGATGAGACATATAGATATATATGAGATGAGAGA 831
 |||||||
 Db 772 CTTATCTCTAAATATATCAAGATGAGACATATAGATATATATGAGATGAGAGA 1017
 |||||||
 QY 832 CATAGTTACAGAGAAATCTAGTCCGAGAGTGAACAGCAAGAGATTTACTGTTACAG 891
 |||||||
 Db 832 CATAGTTACAGAGAAATCTAGTCCGAGAGTGAACAGCAAGAGATTTACTGTTACAG 1077
 |||||||
 QY 892 TATGATGATCAGAAATATGTAAGCGCTTTCAGACAAACAAATCAGATGAGATGAA 951
 |||||||
 Db 892 TATGATGATCAGAAATATGTAAGCGCTTTCAGACAAACAAATCAGATGAGATGAA 1137
 |||||||
 QY 952 AACACATTTGGAATGCAAGCAATTTCTCAGAGCCATACAGTTGAGTCTGCTCCAG 1011
 |||||||
 Db 952 AACACATTTGGAATGCAAGCAATTTCTCAGAGCCATACAGTTGAGTCTGCTCCAG 1197
 |||||||
 QY 1012 TATGATGAGAGAGTATCATTAACAGATCATCGATTCACAGTCCAGATGGATGTA 1071
 |||||||
 Db 1012 TATGATGAGAGAGTATCATTAACAGATCATCGATTCACAGTCCAGATGGATGTA 1257
 |||||||
 QY 1072 AATACAGTGAATATGCTAAACGTTGATCACCATTGGAAGCATTCGACATTCGCT 1131
 |||||||
 Db 1072 AATACAGTGAATATGCTAAACGTTGATCACCATTGGAAGCATTCGACATTCGCT 1317
 |||||||
 QY 1132 TTCAATATGCGATGATGCTGACCTGCTCCAAAGATCGTTCCATTCGTTATGAGATG 1191
 |||||||
 Db 1132 TTCAATATGCGATGATGCTGACCTGCTCCAAAGATCGTTCCATTCGTTATGAGATG 1377
 |||||||
 QY 1192 GCTCTCCCACTGACATTCCTCCGGAGAGTGTGTGCGAGACCGAGCTGCTCAAT 1251
 |||||||
 Db 1192 GCTCTCCCACTGACATTCCTCCGGAGAGTGTGTGCGAGACCGAGCTGCTCAAT 1437
 |||||||
 QY 1252 GTTGTGACCTTTGATGACAGTACATGTTTCTGCATCTGGGAGATGAGACTTAAGGTA 1311
 |||||||
 Db 1252 GTTGTGACCTTTGATGACAGTACATGTTTCTGCATCTGGGAGATGAGACTTAAGGTA 1497
 |||||||
 QY 1312 TGGACACAAATGATCTGTGAATTTGTAGGACCTTAATGACACAAAGAGCATTCGC 1371
 |||||||
 Db 1312 TGGACACAAATGATCTGTGAATTTGTAGGACCTTAATGACACAAAGAGCATTCGC 1557
 |||||||
 QY 1372 TGTTCGATGACAGGAGGAGGCTGAGTGAAGGCTCATCTGACAACTATCAGATTA 1431
 |||||||
 Db 1372 TGTTCGATGACAGGAGGAGGCTGAGTGAAGGCTCATCTGACAACTATCAGATTA 1617
 |||||||
 QY 1432 TGGACATAGAAATGCTGCATGTTTACAGTGTGTAAGAGCCAGAGAAATTCGCT 1491
 |||||||
 Db 1432 TGGACATAGAAATGCTGCATGTTTACAGTGTGTAAGAGCCAGAGAAATTCGCT 1677
 |||||||
 QY 1492 TGTATTCGATTTATTAACAAGAGATATGTCAGTGGGCGCTATGATGGAATAATTAAGTG 1551
 |||||||
 Db 1492 TGTATTCGATTTATTAACAAGAGATATGTCAGTGGGCGCTATGATGGAATAATTAAGTG 1737
 |||||||
 QY 1552 TGGATCTTGTGCTGCTTGGACCCCGTGTCTGACAGGACATCTGTCTACGGACC 1611
 |||||||
 Db 1552 TGGATCTTGTGCTGCTTGGACCCCGTGTCTGACAGGACATCTGTCTACGGACC 1797
 |||||||
 QY 1612 CTTGTGAGAGCATTCGGAAGAGTTTTCGATACAGTGTGTAAGATTCAGATTCAGT 1671
 |||||||
 Db 1612 CTTGTGAGAGCATTCGGAAGAGTTTTCGATACAGTGTGTAAGATTCAGATTCAGT 1857
 |||||||
 QY 1672 AGTTCAATGATGACAAATCTCATCTGGAGCTTCTTAATATGATCCAGTGGCCAAGT 1731
 |||||||
 Db 1672 AGTTCAATGATGACAAATCTCATCTGGAGCTTCTTAATATGATCCAGTGGCCAAGT 1917
 |||||||
 QY 1732 GAAACCCCGCTTCCCTTCGACACATACACCTACATCTCAGATTAATTAACATACAG 1791
 |||||||
 Db 1732 GAAACCCCGCTTCCCTTCGACACATACACCTACATCTCAGATTAATTAACATACAG 1977
 |||||||
 QY 1792 TGACCTCATCTGCTCCAGAGACCCATTTAAG-TTGGCGTATTTAACGTAATCTCCCAATAC 1850
 |||||||

Db 1978 TGACCTCATCTGCTCCAGAGACCCATTAAGTTGGGATTTAAGTATCTGCCAATAC 2037
 |||||||
 QY 1851 CAGATGAGCAACACAGTATACATCAAACTCTGCCAGTTTCCCTGATACAGCCGAG 1910
 |||||||
 Db 2038 CAGATGAGCAACACAGTATACATCAAACTCTGCCAGTTTCCCTGATACAGCCGAG 2097
 |||||||
 QY 1911 AGAGAGGCTTTGAGATCTCTGTTGGAGACAGTTGGTGTGACAGTGGCCAGAGGGTCT 1970
 |||||||
 Db 2098 AGAGAGGCTTTGAGATCTCTGTTGGAGACAGTTGGTGTGACAGTGGCCAGAGGGTCT 2157
 |||||||
 QY 1971 ACTGACACAACTGACTGCTTCAGTCTGCTATACAGAGATGCTTTCATCAATTTGAA 2030
 |||||||
 Db 2158 ACTGACACAACTGACTGCTTCAGTCTGCTATCAAAAGATGCTTTCATCTTTGTGAA 2217
 |||||||
 QY 2031 TGATTTGAACTTTTAAACCTCCCTCTCTCTCTTCAACCTGACCTGAGTTTTC 2090
 |||||||
 Db 2218 TGATTTGAACTTTTAAACCTCCCTCTCTCTCTTCAACCTGACCTGAGTTTTC 2277
 |||||||
 QY 2091 CCAATGCTTCAGACAAAGGTGACTTATAATATATTTAGTGTTCAGCA 2143
 |||||||
 Db 2278 CCAATGCTTCAGACAAAGGTGACTTATAATATATTTAGTGTTCAGCA 2230
 |||||||

RESULT 10
 AAK52700
 .ID AAK52700 standard; cDNA; 3003 BP.
 AC AAK52700;
 AC AAK52700; (first entry)
 DT 06-NOV-2001
 XX Human polynucleotide SEQ ID NO 2229.
 DE Human polynucleotide SEQ ID NO 2229.
 XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.
 XX WC0200157190-A2.
 PN 09-AUG-2001.
 PD 05-FEB-2001; 2001WO-US04098.
 PF 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 19-JUL-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HXSE-) HXSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 DR WPT: 2001-476283/51.
 DR P-PSDB: AAM79567.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 useful in diagnosis and gene therapy -
 PS Claim 1: Page 4585-4586; 6221pp; English.
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce

Oy	2031	TGATTTGGAACTTATAAACCTCCCGTCGTCCACCTTTACCGCTCACCCTTGTTTTC	2090
Dd	2218	TGATTTGGAACTTATAAACCTCCCGTCGTCCACCTTTACCGCTCACCCTTGTTTTC	2277
Oy	2091	CCATTGCTGCAGACAAGAAGTGACTTAATAAATAATTAAGTGTGTTGCCAGAA	2143
Dd	2278	CCATTGCTGCAGACAAGAAGTGACTTAATAAATAATTAAGTGTGTTGCCAGAA	2330
<hr/>			
RESULT 11			
ID	AAK52701	standard; cDNA; 3003 BP.	
XX	AAK52701;		
AC	AAK52701;		
XX	06-NOV-2001	(first entry)	
DE	Human polynucleotide SEQ ID NO 2230.		
XX			
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorder; arthritis; inflammation; ss.		
OS	Homo sapiens.		
XX			
PN	WO200157190-A2.		
PD	09-AUG-2001.		
PE	05-FEB-2001; 2001MO-USO4098.		
PR	03-FEB-2000; 2000US-0496914.		
PR	27-APR-2000; 2000US-0560875.		
PR	20-JUN-2000; 2000US-0598075.		
PR	19-JUL-2000; 2000US-0620325.		
PR	01-SEP-2000; 2000US-0654936.		
PR	15-SEP-2000; 2000US-0663561.		
PR	20-OCT-2000; 2000US-0693325.		
PR	30-NOV-2000; 2000US-0728422.		
XX			
PA	(HYSE-) HYSEQ INC.		
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Gao Y, Ma Y;		
PI	Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		
PL	Xue A, Yang Y, Wejtnan T, Goodrich R;		
XX	WPI: 2001-476283/51.		
DR	P-PsDB: AAM79568.		
PT	Nucleic acids encoding polypeptides with cytokine-like activities,		
PT	useful in diagnosis and gene therapy -		
XX			
PS	Claim 1; Page 4586-4587; 6221pp; English.		
XX	The invention relates to polynucleotides (AAK51456-AAK53435) and the		
CC	encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, hematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		
CC	activin/inhibin activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and		
CC	inflammation.		
CC	Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666		
CC	(AAM80020) are omitted as the relevant pages from the sequence listing		
CC	were missing at the time of publication.		
XX			
S0	Sequence 3003 BP; 807 A; 703 C; 722 G; 772 T; 0 other:		

Query Match 93.4%; Score 2008.2; DB 22; Length 3003;

	Best Local Similarity: 99.6%;	Pred. No. 0;	Mismatches 8;	Indels 1;	Gaps 1;
Matches 2024:	Conservative	0;			
QY	112	TTTTGTAATTCCTCAGAGAGAGAGACTGTAAATTAATGGCAACCCCTTAGAAGATTAATA	171		
	11				
Db	298	TTCCAGAAATTCCTCAGAGAGAGAGAGACTGTAAATTAATGGCAACCCCTTAGAAGATTAATA	357		
QY	172	CCAGAGAAATTCCTCCTTGACAGACATTCACAGCTGTGGCCAGACTGTGGCTTAACCAA	231		
Db	358	CCAGAGAAATTCCTCCTTGACAGACATTCACAGCTGTGGCCAGACTGTGGCTTAACCAA	417		
QY	232	GAACAGTATGTTTAGCAAGCACTGCTATGAAGACTGAGAAATTTGTGTGCCAAAAACAAA	291		
Db	418	GAACAGTATGTTTAGCAAGCACTGCTATGAAGACTGAGAAATTTGTGTGCCAAAAACAAA	477		
QY	292	CTTGGCAATGGACATTCCTCAGTATGATGTGGCCCAAGCAAGGAAACTCTCACCACAGCAT	351		
Db	478	CTTGGCAATGGACATTCCTCAGTATGATGTGGCCCAAGCAAGGAAACTCTCACCACAGCAT	537		
QY	352	GAAGAGGAAAGAAAGAACTGTGTCTCAAAATCTACTTTGAGCAGTGGTTCAGAGTCAGATCAAGTG	411		
Db	538	GAAGAGGAAAGAAAGAACTGTGTCTCAAAATCTACTTTGAGCAGTGGTTCAGAGTCAGATCAAGTG	597		
QY	412	GAAATTTGTGGAAACATCTTATATCCCAATATGTTCATTACCAACATGGGCACATTAACCTCG	471		
Db	598	GAAATTTGTGGAAACATCTTATATCCCAATATGTTCATTACCAACATGGGCACATTAACCTCG	657		
QY	472	TATCTTAAACCTATGTTCGAGAGATTTCAATACGCTCTCCAGACTGGGGATTTGGAT	531		
Db	658	TATCTTAAACCTATGTTCGAGAGATTTCAATACGCTCTCTCCAGACTGGGGATTTGGAT	717		
QY	532	CATATCGCTGAGACATTTCTGTCAATACCTGGATGCCAAATCTATATGCTGCTGAACCT	591		
Db	718	CATATCGCTGAGACATTTCTGTCAATACCTGGATGCCAAATCTATATGCTGCTGAACCT	777		
QY	592	GGTGGCAAGAAATGGATACCGAGTACCTGTATGGCATGCTGTGGAAAGAGCTTATACAG	651		
Db	778	GGTGGCAAGAAATGGATACCGAGTACCTGTATGGCATGCTGTGGAAAGAGCTTATACAG	837		
QY	652	AGAATGGTCAGAGACAGATTCCTGTGGAGAGCCCTGGCAGAAAGAAAGAGATGGGAGACG	711		
Db	838	AGAATGGTCAGAGACAGATTCCTGTGGAGAGCCCTGGCAGAAAGAAAGAGATGGGAGACG	897		
QY	712	TATTTATTCOAACAAACCTCCTCAGAGGGAATGGTCCGCCCAACTCTTTTATATAGAGA	771		
Db	898	TATTTATTCOAACAAACCTCCTCAGAGGGAATGGTCCGCCCAACTCTTTTATATAGAGA	957		
QY	772	CTTTATTCCTAAATTTATACAGACATTTGAGACAAATAGATCTTAATTTGGAGATGGGAAGA	831		
Db	958	CTTTATTCCTAAATTTATACAGACATTTGAGACAAATAGATCTTAATTTGGAGATGGGAAGA	1017		
QY	832	CATAGTTTACAGAGAAATTCACCTCCGGAAGTGAACAACAAGAAAGAGATTACTGTTTACAG	891		
Db	1018	CATAGTTTACAGAGAAATTCACCTCCGGAAGTGAACAACAAGAAAGAGATTACTGTTTACAG	1077		
QY	892	TATGATGATCAAGAAATATGTAAGCGGCTCTCAGACAACACATATAGATCTGGGATAAA	951		
Db	1078	TATGATGATCAAGAAATATGTAAGCGGCTCTCAGACAACACATATAGATCTGGGATAAA	1137		
QY	952	AACACATTTGGAATGCAAGCGAAATTCACAGGCCATACACAGTTTCAAGTCTGTCTCCAG	1011		
Db	1138	AACACATTTGGAATGCAAGCGAAATTCACAGGCCATACACAGTTTCAAGTCTGTCTCCAG	1197		
QY	1012	TATGATGAGAGAGTATCATTAACAGAGTATGAGATTCACAGCTGTAGAGATGGGATGTA	1071		
Db	1198	TATGATGAGAGAGTATCATTAACAGAGTATGAGATTCACAGCTGTAGAGATGGGATGTA	1257		
QY	1072	AATACAGGTGAATGCTAAMACGTTTGATTCACATTTGTGAAGCAGTTCTGCACATCTGGGT	1131		
Db	1258	AATACAGGTGAATGCTAAMACGTTTGATTCACATTTGTGAAGCAGTTCTGCACATCTGGGT	1317		
QY	1132	TTTCATTAATGCGATGATGGTGACCTGTCTCAAAAGATCGTTCATTGCTGTATGGGATATG	1191		
	11				


```

Db 361 ATATCCAAATGTGTCATTTACCAACATGSGGACATMAACTGTAATCTTAAACTATGTTG 420
Oy 490 CAGAGAGATTTTCATTAAGTCTGTCGACGTCGGGAGTTGATTCATATGCTGAGAACAT 549
Db 421 CAGAGAGATTTTCATTAAGTCTGTCGACGTCGGGAGTTGATTCATATGCTGAGAACAT 480
Oy 550 CTGTATACCTTGGATGCGCAAAATCACTATGCTGCTGAACTTGTGTGCAAGAAATGATC 609
Db 481 CTGTATACCTTGGATGCGCAAAATCACTATGCTGCTGAACTTGTGTGCAAGAAATGATC 540
Oy 610 CGAGTACCTTGGATGCGCAAAATCACTATGCTGCTGAACTTGTGTGCAAGAAATGATC 669
Db 541 CGAGTACCTTGGATGCGCAAAATCACTATGCTGCTGAACTTGTGTGCAAGAAATGATC 600
Oy 670 TCTGTGTGAGAGGCGCTGGGACAAAGAGAGATGGGAGACATTTATTTGAAAAACAAA 729
Db 601 TCTGTGTGAGAGGCGCTGGGACAAAGAGAGATGGGAGACATTTATTTTAAAAACAAA 660
Oy 730 CCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACTTTATCTTAAATTTATA 789
Db 661 CCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACTTTATCTTAAATTTATA 720
Oy 790 CAAGCATTTGAGACATTAAGATCTAATTTGAGATGTTGAGAGATAGTTTACAGAAAT 849
Db 721 CAAGCATTTGAGACATTAAGATCTAATTTGAGATGTTGAGAGATAGTTTACAGAAAT 780
Oy 850 CACTGCGGAGTTGAAACAAAGAGTTTACTCTTTTACAGTATGATGATCAGAAATA 909
Db 781 CACTGCGGAGTTGAAACAAAGAGTTTACTCTTTTACAGTATGATGATCAGAAATA 840
Oy 910 GTAAGGCGGCTTCGAGACAAACAAATCAAGATCTGGATTAACAAACATTTGGAAATGCA 969
Db 841 GTAAGGCGGCTTCGAGACAAACAAATCAAGATCTGGATTAACAAACATTTGGAAATGCA 900
Oy 970 CGAATTTTCACAGGCGCATACAGGTTCACTCTCTCTCAGTATGATGAGAGATGATC 1029
Db 901 CGAATTTTCACAGGCGCATACAGGTTCACTCTCTCTCAGTATGATGAGAGATGATC 960
Oy 1030 ATAACAGATCATCGGATTCACAGGTCAGAGTGTGGATGTAAATACAGGTCGAAATGCTA 1089
Db 961 ATAACAGATCATCGGATTCACAGGTCAGAGTGTGGATGTAAATACAGGTCGAAATGCTA 1020
Oy 1090 AACAGCTTATTCACATGTTGAGACAGTGTGCACTGCTGCTTCAATTAATGCGATGATG 1149
Db 1021 AACAGCTTATTCACATGTTGAGACAGTGTGCACTGCTGCTTCAATTAATGCGATGATG 1080
Oy 1150 GTGACCTGCTCCAAAGATGCTTCATTTGCTGTATGGATATGCGCTCCCAACTGACAT 1209
Db 1081 GTGACCTGCTCCAAAGATGCTTCATTTGCTGTATGGATATGCGCTCCCAACTGACAT 1140
Oy 1210 ACCCTCGGAGGCTGCTGCTGCGACACGAGTGTGCTCAATGTTGAGCTTGTATGAC 1269
Db 1141 ACCCTCGGAGGCTGCTGCTGCGACACGAGTGTGCTCAATGTTGAGCTTGTATGAC 1200
Oy 1270 AAGTACATTTGTTTCTGATCTGAGGATAGACATTAAGGATATGAGACAAAGTACTGT 1329
Db 1201 AAGTACATTTGTTTCTGATCTGAGGATAGACATTAAGGATATGAGACAAAGTACTGT 1260
Oy 1330 GAATTTGTAAAGGACCTTAATATGAGACAAAGAGCATGCTGTTTGGAGTACAGGAC 1389
Db 1261 GAATTTGTAAAGGACCTTAATATGAGACAAAGAGCATGCTGTTTGGAGTACAGGAC 1320
Oy 1390 AGGCTGTAGTGTGCTGCTGCTGCAACACTATCAGTTATGGACATAGAAATGTGT 1449
Db 1321 AGGCTGTAGTGTGCTGCTGCTGCAACACTATCAGTTATGGACATAGAAATGTGT 1380
Oy 1450 GCATTTGTAAAGGCTTAAAGGACCTGAGAAATGCTGCTGTTGATGATTTGATTAAC 1509
Db 1381 GCATTTGTAAAGGCTTAAAGGACCTGAGAAATGCTGCTGTTGATGATTTGATTAAC 1440
Oy 1510 AAGAGATATGCTAGTGGGCGCTATGATGAAAAATTTAAAGTGTGGATCTTGTGCTGT 1569
Db 1441 AAGAGATATGCTAGTGGGCGCTATGATGAAAAATTTAAAGTGTGGATCTTGTGCTGT 1500

```

```

Oy 1570 TTGAGCCCCGCTGCTCTCGAGAGGACACTCTGTCTACGAGACCTTGTGAGCATTCGGA 1629
Db 1501 TTGAGCCCCGCTGCTCTCGAGAGGACACTCTGTCTACGAGACCTTGTGAGCATTCGGA 1560
Oy 1630 AGAGTTTGTGACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1689
Db 1561 AGAGTTTGTGACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Oy 1690 ATCCCTATCTGGGACCTTCTAAATGATTCAGCTGCCCCAAGCTGAGACCCCCGTTCCCT 1749
Db 1621 ATCCCTATCTGGGACCTTCTAAATGATTCAGCTGCCCCAAGCTGAGACCCCCGTTCCCT 1680
Oy 1750 TCTGACATATACACTTACATCTCCAGA 1776
Db 1681 TCTGACATATACACTTACATCTCCAGA 1707

RESULT 13
AAK51715
ID AAK51715 standard; cdna; 2285 bp.
XX
XX AAK51715;
AC
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 260.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
XX MO200157190-A2.
XX
PD 09-AUG-2001.
XX
XX 05-FEB-2001; 2001MO-0504098.
XX
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YF, Liu C, Dymnac RF, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AD, Yang Y, Wejrtman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
XX
XX P-PSDB; AAM78582.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX Claim 1; Page 1177-1180; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and/or
CC activin/inhibin activity and may be useful in the diagnosis and/or

```


OY 1552 TGGATCTGTCGCTGCTTGGACCCCGCTCTCTGACGAGACATCTGTCTACGAGC 1611
 DB 1798 TGGATCTGTCGCTGCTTGGACCCCGCTCTCTGACGAGACATCTGTCTACGAGC 1857
 OY 1612 CTGTGAGCATCTCGGAGAGAGTTTTCGACTACAGTTTGAATTCAGATTCAGT 1671
 DB 1858 CTGTGAGCATCTCGGAGAGAGTTTTCGACTACAGTTTGAATTCAGATTCAGT 1917
 OY 1672 AGTTACATGATGACACAAATCTCTATCTGAGACTTCTCTAAATGATCCAGCTGCAAGCT 1731
 DB 1918 AGTTACATGATGACACAAATCTCTATCTGAGACTTCTCTAAATGATCCAGCTGCAAGCT 1977
 OY 1732 GAACCCCGCTGCTCTCTCTGACATACACCTACATCTCTGAGATTAATACCATAC 1791
 DB 1978 GAACCCCGCTGCTCTCTCTGACATACACCTACATCTCTGAGATTAATACCATAC 2037
 OY 1792 TGACCTCATCTGCGCAGGACCATTAAGTTGCGGTATTAAAC 1836
 DB 2038 TGACCTCATCTGCGCAGGACCATTAAGTTGCGGTATTAAAC 2082

RESULT 15

AKS1716 standard; cDNA; 2207 BP.

AKS1716;

06-NOV-2001 (first entry)

Human polynucleotide SEQ ID NO 261.

Human: cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KM tissue growth factor; immunomodulatory; cancer; leukemia;

OS nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.

XX MO200157190-A2.

XX 09-ANG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX (HSE-) HSEQ INC.

PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

DR WPI: 2001-476283/51.

DR P-PSDB: AAM78583.

PT Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

PS Claim 1; Page 1180-1183; 6221pp; English.

CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AKS2581), 2111 (AKS2582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

Sequence 2207 BP; 615 A; 478 C; 555 G; 559 T; 0 other:

Query Match 78.5%; Score 1688.8; DB 22; Length 2207;
 Best Local Similarity 94.8%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 67; Indels 30; Gaps 1;

OY 1 TGGGTGGCTGGGCTGCTGACCAAAAGGGGCGGCGGAGAGCCAGTGGCC 60
 DB 58 TGGGTGGCTGGGCTGCTGACCAAAAGGGGCGGCGGAGAGCCAGTGGCC 117
 OY 61 TGGGCGATTATGAGCCCGCGGCGGCTGCTGCAAGAGAGCACTCAAGTTATGAT 120
 DB 118 TGGGCGATTATGAGCCCGCGGCGGCTGCTGCAAGAGAGCACTCAAGTTATGAT 177
 OY 121 TCGTCAAGAGAGAAAGACTGTAAATATGCGCAACCCCTAGAGATTAATCCAGAGAG 180
 DB 178 TCGTCAAGAGAGAAAGACTGTAAATATGCGCTGCTGCTGCGGAGAGATGCGCTGCTG 237
 OY 181 AATT-----CACTTAGACAGACTACACAGCTGT 210
 DB 238 CGATGCTCTTAAATACCCAGGAGCTGCGCACTCAAGCTTTTCCAGATTAACAGCTGT 297
 OY 211 GCGAGACTCTGCTTAAACCAAGAGAGAGTATGAGAGAGAGAGAGAGAGAGAGAGAG 270
 DB 298 GCGAGACTCTGCTTAAACCAAGAGAGAGTATGAGAGAGAGAGAGAGAGAGAGAGAG 357
 OY 271 AATTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 330
 DB 358 AATTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 417
 OY 331 CGGAAGCTCTGCAAG 390
 DB 418 CGGAAGCTCTGCAAG 477
 OY 391 TGGTCAAG 450
 DB 478 TGGTCAAG 537
 OY 451 CAACATGGGCGAG 510
 DB 538 CAACATGGGCGAG 597
 OY 511 CTGCCAGCTGGGCGGAG 570
 DB 598 CTGCCAGCTGGGCGGAG 657
 OY 571 TCACATGTGCTGCTGCAAG 630
 DB 658 TCACATGTGCTGCTGCAAG 717
 OY 631 CTGTGAG 690
 DB 718 CTGTGAG 777
 OY 691 GAAG 750
 DB 778 GAAG 837
 OY 751 CCCAGCTCTTTTATAG 810
 DB 838 CCCAGCTCTTTTATAG 897
 OY 811 TCTAATGAG 870

Db 898 TCTAATTGGAGATGTGGAAACATAGTTACAGAGATTCACTGCCGAAGTGAACACAGC 957
QY 871 AAGGAGTTTACTGTTTACAGTATGATGATCAGAAATAGTAAGCGCCCTTCGAGACAC 930
Db 958 AAGGAGTTTACTGTTTACAGTATGATGATCAGAAATAGTAAGCGCCCTTCGAGACAC 1017
QY 931 ACAATCAAGATCTGGGATAAAAACATATTGGAATGCAAGCAATTTCTCAGAGCCATACA 990
Db 1018 ACAATCAAGATCTGGGATAAAAACATATTGGAATGCAAGCAATTTCTCAGAGCCATACA 1077
QY 991 GGTTCAGTCTCTGTTCCAGTATGATGAGAGAGATGATCATTAACAGATCATCGATTCC 1050
Db 1078 GGTTCAGTCTCTGTTCCAGTATGATGAGAGAGATGATCATTAACAGATCATCGATTCC 1137
QY 1051 ACGGTACAGTGTGGAGTGAATACAGGTGAATGCTAAACAGTTCATCCATTGT 1110
Db 1138 ACGGTACAGTGTGGAGTGAATACAGGTGAATGCTAAACAGTTCATCCATTGT 1197
QY 1111 GAAGCATTTCTGACTTCGTTCCATTAATGCGATGATGTTGCTCCAAAGATCGT 1170
Db 1198 GAAGCATTTCTGACTTCGTTCCATTAATGCGATGATGTTGCTCCAAAGATCGT 1257
QY 1171 TCCATTGCTGTATGGATATGGCTCCCAACATGACATTAACCTCCGAGGGTGTGGTC 1230
Db 1258 TCCATTGCTGTATGGATATGGCTCCCAACATGACATTAACCTCCGAGGGTGTGGTC 1317
QY 1231 GGACACCGGAGCTGCTGTCATTTGTTAGACTTTGATGACACAGTACATTTGTTCTGCATCT 1290
Db 1318 GGACACCGGAGCTGCTGTCATTTGTTAGACTTTGATGACACAGTACATTTGTTCTGCATCT 1377
QY 1291 GGGGATGAACTATTAAGATATGGAACACAAAGTACTTGTGAATTTGTAAGACCTTAAT 1350
Db 1378 GGGGATGAACTATTAAGATATGGAACACAAAGTACTTGTGAATTTGTAAGACCTTAAT 1437
QY 1351 GGACACAACGAGGCAATGGCTGTTGTCAGTACAGGACAGGCTGGTACTGAGTGGCTCA 1410
Db 1438 GGACACAACGAGGCAATGGCTGTTGTCAGTACAGGACAGGCTGGTACTGAGTGGCTCA 1497
QY 1411 TCTGACAACACTATACAGATTATGGACATAGGAATGTTGTCAGTATGAGTGTAGAA 1470
Db 1498 TCTGACAACACTATACAGATTATGGACATAGGAATGTTGTCAGTATGAGTGTAGAA 1557
QY 1471 GGCCATGAGGAATTTGTCGTTGATTGATTGATTGAACAAGAGATGACAGTGGGGCC 1530
Db 1558 GGCCATGAGGAATTTGTCGTTGATTGATTGATTGAACAAGAGATGACAGTGGGGCC 1617
QY 1531 TATGATGAAAAATTTAAAGTGTGGGATCTTGTGCTCTTTGGACCCCGCTGCTGCA 1590
Db 1618 TATGATGAAAAATTTAAAGTGTGGGATCTTGTGCTCTTTGGACCCCGCTGCTGCA 1677
QY 1591 GGGACACTCTGTACGAGCCCTTGTGAGCATTCGGAAGAGTTTTCGACTACAGTTT 1650
Db 1678 GGGACACTCTGTACGAGCCCTTGTGAGCATTCGGAAGAGTTTTCGACTACAGTTT 1737
QY 1651 GATGAATTCAGATTGTCAGTAGTTCAATGATGACACAATCTCATCTGGGACTTCTTA 1710
Db 1738 GATGAATTCAGATTGTCAGTAGTTCAATGATGACACAATCTCATCTGGGACTTCTTA 1797
QY 1711 AATGATCCAGCTGGCCCAAGTGAACCCCGCTTCCCTTCTGGAACATACACTACATC 1770
Db 1798 AATGATCCAGCTGGCCCAAGTGAACCCCGCTTCCCTTCTGGAACATACACTACATC 1857
QY 1771 TCCAGATTAATTAACCATACACTGACCTCATATTGCCAGGACCCATTAAGTTGCGGTA 1830
Db 1858 TCCAGATTAATTAACCATACACTGACCTCATATTGCCAGGATTCGAATCGATTATGTA 1917
QY 1831 TTTAAC 1836
Db 1918 CATTAAC 1923

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 22:02:00 : Search time 2778 Seconds

(without alignments)
12540.134 Million cell updates/sec

Title: US-09-601-168b-1

Perfect score: 2151

Sequence: 1 tgcgttgctgcgcctggc.....gtttgcagaaaaaaaaa 2151

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlinu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlinu:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	650.6	30.2	709	12	BE782628 601465751
2	638	29.7	778	14	BE782628 601465751
3	635.2	29.5	652	9	AL044095
4	625	29.1	851	13	AL044095 DKFZP434M
5	624.8	29.0	973	14	BM465280
6	613.2	28.5	758	10	BE296484

7	608.8	28.3	692	10	BE269297
8	589.4	27.4	826	13	BE22845
9	586.6	27.3	1147	14	BQ948186
10	580.6	27.0	743	12	BG829037
11	566.4	26.3	752	12	BG722472
12	534.4	24.8	789	12	BG518761
13	532.4	24.8	637	13	BI103092
14	529.8	24.6	544	9	AL041438
15	528.4	24.6	888	13	BI455345
16	516.2	24.0	810	14	BM944304
17	500.4	23.3	646	12	BG082149
18	497	23.1	850	14	BQ771102
19	492.2	22.9	541	12	BG058605
20	488.8	22.7	590	10	BE762522
21	483.8	22.5	521	12	BF076123
22	483.8	22.5	945	9	AL556068
23	483.6	22.5	580	12	BG277740
24	472.8	22.0	732	13	BI088274
25	471.4	21.9	658	12	BF781002
26	468	21.8	656	9	AL707470
27	464.8	21.6	972	14	BQ940482
28	459.2	21.3	665	9	AL642623
29	458.8	21.3	704	13	BQ48898
30	456.6	21.2	695	13	BM214068
31	452.8	21.1	682	10	BE109603
32	445.6	20.7	666	14	BQ211349
33	445.2	20.7	553	9	AA478504
34	444.4	20.6	717	10	AV716543
35	443.6	20.6	577	10	BE284066
36	440	20.5	621	14	BQ383756
37	438.8	20.4	506	9	AI430490
38	438.2	20.4	509	14	BQ340149
39	425.6	19.8	664	13	BM211073
40	419.6	19.5	642	13	BI278543
41	417.6	19.4	656	13	BM214360
42	416.2	19.3	1081	13	BM542586
43	414.8	19.3	524	9	AA033076
44	405.4	18.8	538	9	AA230588
45	404.2	18.8			

ALIGNMENTS

RESULT 1
LOCUS BE782628
DEFINITION 601465751F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3868812 5', mRNA sequence.
VERSION BE782628.1 GI:10203826
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 709)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1989)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM9617 row: f column: 13
High quality sequence stop: 655.
location/Qualifiers
I..709

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3868812"
 /clone_1id="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SORI6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

BASE COUNT 193 a 156 c 175 g 185 t
 ORIGIN

Query Match 30.2%; Score 650.6; DB 12; Length 709;
 Best Local Similarity 98.2%; Pred. No. 7.5e-184;
 Matches 700; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

1222 GTGCTGTCGACACGAGCTGCTCAATGTTGTAGACTTTGATGACAGTACATTTGT 1281
 1 GTGCTGTCGACACGAGCTGCTCAATGTTGTAGACTTTGATGACAGTACATTTGT 59
 1282 TCTGATCTGGGATAGACTATTAAGTATGAAACACAGTACTTGTGAATTTGTAAG 1341
 60 TCTGATCTGGGATAGACTATTAAGTATGAAACACAGTACTTGTGAATTTGTAAG 119
 1342 ACCATAATGACACAAACGAGGATGCTGTTGCTGAGTACAGGAGGAGGCTGTAGT 1401
 120 ACCATAATGACACAAACGAGGATGCTGTTGCTGAGTACAGGAGGAGGCTGTAGT 179
 1402 AGTGCGTCATGTACACACTATGATGATGAGATGAGATGATGATGATGATGATGAT 1461
 180 AGTGCGTCATGTACACACTATGATGATGAGATGAGATGATGATGATGATGATGATGAT 239
 1462 GTGTAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1521
 240 GTGTAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299
 1522 AGTGCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1581
 300 AGTGCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
 1582 GCTCTGAGGAGGACACTGCTGCTACGACGCTTGTGAGGACATTCGGAAGAGTTTTCGA 1641
 360 GCTCTGAGGAGGACACTGCTGCTACGACGCTTGTGAGGACATTCGGAAGAGTTTTCGA 419
 1642 CTACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1701
 420 CTACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
 1702 GACTTCTCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1761
 480 GACTTCTCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
 1762 ACCATCATCTCCAGATTAATTAACATACATGATGATGATGATGATGATGATGATGAT 1821
 539 ACCATCATCTCCAGATTAATTAACATACATGATGATGATGATGATGATGATGATGAT 598
 1822 GTTGGGATTTTAACGATCTGCAATACAGATGAGCAACAAAGTAAACATCAAACT 1881
 599 GTTGGGATTTTAACGATCTGCAATACAGATGAGCAACAAAGTAAACATCAAACT 656
 1882 ACTGGCCAGTTTCCCTGACACTAGCCGAGGAGGAGGCTTTGAGAGCTCTGTG 1934
 657 AGTGGCCAGTTTCCCTGACACTAGCCGAGGAGGAGGCTTTGAGAGCTCTGTG 709

RESULT 2
 LOCUS B0444188 778 bp mRNA linear EST 29-MAY-2002
 DEFINITION U1-M-EX0-bx1-1-06-0-UT.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
 IMAGE:5709077 5, mRNA sequence.
 ACCESSION B0444188
 VERSION B0444188.1 GI:21247300

KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 778)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLN at:
<http://image.jiml.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

source

Location/Qualifiers
 1..778

/organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5709077"
 /clone_1id="NIH_BMAP_EX0"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /note="Organ: brain; Vector: pX-asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pX-asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GTGCGTGGA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 223 a 181 c 204 g 165 t 1 others

Query Match 29.7%; Score 638; DB 14; Length 778;
 Best Local Similarity 89.5%; Pred. No. 4.9e-180;
 Matches 697; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

256 GCTATGAAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315
 1 GCTATGAAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 316 ATGTGCCAAGAGGAGGAACTCTCAGACAGTATGAAAGGAAAGAACTGTGTGTC 375
 61 ATGTGCCAAGAGGAGGAACTCTCAGACAGTATGAAAGGAAAGAACTGTGTGTC 120
 376 AATATCTTTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
 121 AATATCTTTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 436 CAAATGTCATTAACCAACATGAGGACATAAAGTATCTTAACCTATGTCAGAGA 495
 181 CAAATGTCATTAACCAACATGAGGACATAAAGTATCTTAACCTATGTCAGAGA 240
 496 GATTTCATACGCTGCTGACAGCTGGGAGTGGATCATATGATGATGATGATGATGATGAT 555
 241 GATTTCATACGCTGCTGACAGCTGGGAGTGGATCATATGATGATGATGATGATGATGAT 300

OY 556 TACCGGATGCCAATACATATGCTGCTGCACTTGTGTGCAAGGATGTTACCGAGTG 615
 DB 301 TACTTGGACGCAAGTACATGCTGTGCTGAGCTCGTGGCAAGGATGATGACCGCTG 360
 OY 616 ACTCTGTGATGATGCTGTGTGGAAGACCTTATCGAGAGAAATGTCAGACAGATTCCTG 675
 DB 361 ACCTGCGAGCGCATGCTGTGTGAAAAACCTATCGACAGGATGTCAGACAGGACTCTG 420
 OY 676 TGAAGAGCCCTGGCAGAACGAGATGGGACATATTATTCAAAAACCACTCT 735
 DB 421 TGGGAGAGCCCTGGCAGAGCGAGGCTGGGAGACATATTATTCAAAAACCACTCT 480
 OY 736 GAGCGGATGCTGCTGCCAATCTCTTTTATAGACATTTATCTAAATATTACAAG 795
 DB 481 GATGAGAACGCTCTCTCCCACTCTTTTATAGAGCGCTTATCTTAAATATTACAAG 540
 OY 796 ATTGAGCAATAGATCTAATTTGAGATGTGAAGACATAGTTTACAGACATTTCTG 855
 DB 541 ATTGAGCAATAGATCTAATTTGAGATGTGAAGACATAGTTTACAGACATTTCTG 600
 OY 856 CGAAGTAAACAGCAAGAGATTTACTGTTTACAGATGATGATGAGAAATAGTAGC 915
 DB 601 CGGAGTAAACAGTAAAGGAGTTTACTGTTTACAGATGATGATGAGAAATAGTAGC 660
 OY 916 GGCCTTGAGACAAACATCAAGATCTGGGATTAATAACACATTTGGAATGCAAGCAAT 975
 DB 661 GGCCTTGAGACACACACATCAAGATCTGGGATTAATAACACATTTGGAATGCAAGCAAT 719
 OY 976 CTCACAGCCCATACAGATGCTGCTGCTGCTGCTGATGATGAGAGATGATCATAC 1034
 DB 720 CTCACAGCCCATACAGATGCTGCTGCTGCTGCTGATGATGAGAGATGATCATAC 778

 RESULT 3
 AL044095 652 bp mRNA linear EST 29-FEB-2000
 LOCUS DKFP434M1528_r1 434 (synonym: htes3) Homo sapiens cDNA clone
 DEFINITION DKFP434M1528 5', mRNA sequence.
 ACCESSION AL044095
 VERSION AL044095.1 GI:5432322
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 652)
 Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
 Wiemann, S.
 EST (Bioecker, et al.)
 Unpublished (1999)
 CONTACT: Bioecker H
 MIPS
 Am Klopferstutz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No sl sequence available.
 This clone (DKFP434M1528) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 Location/Qualifiers
 1..652
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFP434M1528"
 /clone_1db="434 (synonym: htes3)"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 165 a 166 c 138 g 183 t
 ORIGIN
 Query Match 29.5%; Score 635.2; DB 9; Length 652;
 Best Local Similarity 99.2%; Pred. No. 3.1e-179;
 Matches 649; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

 OY 1479 GGAATGTGCGCTGTGATTCGATTTGATTAACAAGAGATGACATGGGCGCTATGATGG 1538
 DB 1 GGAATGTGCGCTGTGATTCGATTTGATTAACAAGAGATGACATGGGCGCTATGATGG 60
 OY 1539 AAAAATTAAGTGTGAGATCTTGTGCTCTTTGACCCCTGCTCGCAGGACACT 1598
 DB 61 AAAAATTAAGTGTGAGATCTTGTGCTCTTTGACCCCTGCTCGCAGGACACT 120
 OY 1599 CTGTCTACGAGCCCTGTGTGAGACATTCGGAAGAGTTTTCGATCAAGTTGATGAAT 1658
 DB 121 CTGTCTACGAGCCCTGTGTGAGACATTCGGAAGAGTTTTCGATCAAGTTGATGAAT 180
 OY 1659 CCAGATTTGACAGTTCATGATGACACATCCCATCTGGGACTTCCTAAATGATCC 1718
 DB 181 CCAGATTTGACAGTTCATGATGACACATCCCATCTGGGACTTCCTAAATGATCC 238
 OY 1719 AGCTGCCAAGCTGAACCCCGCTCCCTCTGGAACATACACCTCATCTCCAGATA 1778
 DB 239 AGCTGCCAAGCTGAACCCCGCTCCCTCTGGAACATACACCTCATCTCCAGATA 298
 OY 1779 AATTAACATACACTGACCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1838
 DB 299 AATTAACATACACTGACCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
 OY 1839 ATCTGCCAATACAGAGATGAGCAGACAGATTAACATTAACATTAACATTAACATTA 1898
 DB 359 ATCTGCCAATACAGAGATGAGCAGACAGATTAACATTAACATTAACATTAACATTA 418
 OY 1899 GACTAGCCGAGAGAGAGGCTTTGAGACTCTGTTGGACACAGTTGGTGTGACGTGGC 1958
 DB 419 GACTAGCCGAGAGAGAGGCTTTGAGACTCTGTTGGACACAGTTGGTGTGACGTGGC 478
 OY 1959 CCAGAGCGGTCTACGACAGACAACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2018
 DB 479 CCAGAGCGGTCTACGACAGACAACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
 OY 2019 ATCAATTTGTAATGATGGAATTTAAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCA 2078
 DB 539 ATCTTTGTAATGATGGAATTTAAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCA 598
 OY 2079 CCTAGTTTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2132
 DB 599 CCTAGTTTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652

 RESULT 4
 BQ425358 851 bp mRNA linear EST 23-MAY-2002
 LOCUS AGNCCOURT_7907367 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154838
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ425358
 VERSION BQ425358.1 GI:21120673
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 851)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-rtmail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM13496 row: 1 column: 15
 High quality sequence stop: 697.

FEATURES

Source

Location/Qualifiers
 1..851
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6154838"
 /clone.lib="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: PCMV-SORT6; Site: 1; NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

BASE COUNT 243 a 196 c 223 g 188 t 1 others

ORIGIN

Query Match

Best Local Similarity 29.1%; Score 625; DB 14; Length 851;
 Best Local Similarity 99.1%; Pred. No. 4.2e-176;
 Matches 639; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

FEATURES

source

DEFINITION AGENCOURT_6427660 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502075
 5', mRNA sequence.
 ACCESSION BM465280
 VERSION BM465280.1 GI:18514312
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 973)
 AUTHORS NIH-MGC <http://img.mcl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabds-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM12139 row: k column: 04
 High quality sequence stop: 601.
 Location/Qualifiers
 1..973
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5502075"
 /clone.lib="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: PCMV-SORT6; Site: 1; NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

FEATURES

source

BASE COUNT 282 a 200 c 245 g 246 t
 ORIGIN
 Query Match 29.0%; Score 624.8; DB 13; Length 973;
 Best Local Similarity 97.2%; Pred. No. 5.2e-176;
 Matches 690; Conservative 0; Mismatches 12; Indels 8; Gaps 5;

QY 112 TTTATGAAATTCCTCGACAGAGAAACGCTATATATGGCGAACCCCTAGGACATATA 171
 DB 207 TTCCGAAATTCCTCGACAGAGAAACGCTATATATGGCGAACCCCTAGGACATATA 266
 QY 172 CCAGAGAGAAATTCCTAGACAGACATACAAACAGCTGTGCAGACTGTGCTTAAACCA 231
 DB 267 CCAGAGAGAAATTCCTAGACAGACATACAAACAGCTGTGCAGACTGTGCTTAAACCA 326
 QY 232 GAAACAGATGTTTGGCAAGCACTGCTTGAAGACTGTGCTGCTGCTGCTGCTGCTGCT 291
 DB 327 GAAACAGATGTTTGGCAAGCACTGCTTGAAGACTGTGCTGCTGCTGCTGCTGCTGCT 386
 QY 292 CTGGCAATGGCACTTCAGATATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
 DB 387 CTGGCAATGGCACTTCAGATATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446
 QY 352 GAAAGAGAAAGAACTGTGTCTCAAAATCTTTGAGAGCTGTGCTGCTGCTGCTGCTGCT 411
 DB 447 GAAAGAGAAAGAACTGTGTCTCAAAATCTTTGAGAGCTGTGCTGCTGCTGCTGCTGCT 506
 QY 412 GAATTTGTGACATCTTATATCCCAATGTGCTTATACCAACATGGGACATATAAATCG 471
 DB 507 GAATTTGTGACATCTTATATCCCAATGTGCTTATACCAACATGGGACATATAAATCG 566
 QY 472 TATCTTAAACCTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
 DB 567 TATCTTAAACCTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 626
 QY 532 CATATGCTGAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
 DB 627 CATATGCTGAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
 QY 592 GTGTGCAAGAAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
 DB 687 GTGTGCAAGAAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
 QY 652 ACAAATGTGACAGATTTCTGTGTGAGAGAGCTTGGCAAGAAAGAGATGGGG-ACA 710
 DB 747 ACAAATGTGACAGATTTCTGTGTGAGAGAGCTTGGCAAGAAAGAGATGGGGNACA 806
 QY 711 GATTTATTCACAAACAACTCTGACGCGCAATGCTCTCCCA 755
 DB 807 GATTTATTCACAAACAACTCTGACGCGCAATGCTCTCCCA 851

RESULT 5
 BM465280
 LOCUS BM465280 973 bp mRNA linear EST 05-FEB-2002

QY 718 TTTCAAAAACAACTCTGACGCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
 DB 80 TTTCAAAAACAACTCTGACGCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 139
 QY 778 CTTAAATTTATACAGACATTTGACATATGAAATTTGAGATGTGAAACATTA 837
 DB 140 CTTAAATTTATACAGACATTTGACATATGAAATTTGAGATGTGAAACATTA 199
 QY 838 TTTACAGAAATTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
 DB 200 TTTACAGAAATTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
 QY 898 GATCAGAAATTTAGTAAGCGGCTTTCGAGACACACATCAAGATCTGGGATTAACAC 957
 DB 260 GATCAGAAATTTAGTAAGCGGCTTTCGAGACACACATCAAGATCTGGGATTAACAC 319
 QY 958 TTGGAATGCAAGCAATTTCTACAGGCGCATACAGTTCACTGCTGCTGCTGCTGCTGCT 1017
 DB 320 TTGGAATGCAAGCAATTTCTACAGGCGCATACAGTTCACTGCTGCTGCTGCTGCTGCT 379
 QY 1018 GAGAGATGATCAATACAGATCATGAGATTTCCACGCTGAGATGGGATGTAATTA 1077
 DB 380 GAGAGATGATCAATACAGATCATGAGATTTCCACGCTGAGATGGGATGTAATTA 439
 QY 1078 GGTGAATGCTAAACAGCTTGAATTCACATTTGAAGCAATGCTGCTGCTGCTGCTGCT 1137

Db	440	GGTAAATGCTAAACACGCTGATTCCACCATTTGGAAAGCACTTCTGGACTTTGGCTTTCAAT	499
Qy	1138	AATGGCATGATGTCGACCTGCTCCAAAGATCGTTCCATTGCTGTATGAGATATGCGCTCC	119
Db	500	AATGGCATGATGTCGACCTGCTCCAAAGATCGTTCCATTGCTGTATGAGATATGCGCTCC	559
Qy	1198	CCAAGTCATTTACCCCTCCGAGGCGTCTGTGTGGAAAC-GAGCTGCTGTCAATGTGT	125
Db	560	CCAAGTCATTTACCCCTCCGAGGCGTCTGTGTGGAAACCGGAGCTGTCTCAATGTGT	619
Qy	1257	AGACTTGTGATGACAAAGTACATTTGTTCTGCATCTGGGG--ATGAACTATAAAGTATG	131
Db	620	AGACTTGTGATGACAAAGTACATTTGTTCTGCATCTGGGGCATGAAACTATAAAGTATG	679
Qy	1314	GAAAC-CAGAGTACTTGGAAATTTT---AAGGACCTTAATGGACACAA	1359
Db	680	GAAACCCAAAGTACTTGGAAATTTTGTGTAAGAGACCCCTTAATGGACACAA	729

RESULT 6	BE296484	758 bp	mRNA	linear	EST 20-JUL-2000
LOCUS	BE296484				
DEFINITION	BE296484	601174756p11 NIH_MGC_17	Homo sapiens	CDNA clone IMAGE:3530341 5',	
ACCESSION	BE296484			mRNA sequence.	
VERSION	BE296484.1	GI:9180046			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

BASE COUNT	212 a	149 c	202 g	195 t
ORIGIN				

Query Match	28.5%;	Score 613.2;	DB 10;	Length 758;
Best Local Similarity	96.3%;	Pred. No. 1.4e-172;		
Matches 649;	Conservative 0;	Mismatches 23;	Indels 2;	Gaps 2

0y 602 AATGTCACCGAGTAACTCTGTATGGCATGCTGTGGAAAGACTTATCGAGAGAAATGCTCA 661
Db 1 AATGTTACCGAGTAACTCTGTATGGCATGCTGTGGAAAGACTTATCGAGAGAAATGCTCA 60

0y 662 GGACAGATTCTCTGTGGAGAGCCCTGGCAGAACCAAGATGGGGCACAATAATTATTCA 721

Db	61	GGACGATTTCTCTGTGGAGAGGCGCTGGCAGAAAGAGATGGGACGATTTATTTC	120
Oy	722	AAAACAAACCTTCCTACGGGAATGCTCCTCCCACTCTTTTATAGACACTTATCCTA	781
Db	121	AAAACAAACCTCTGTACGGGAATGCTCCTCCCACTCTTTTATAGACACTTATTCCTA	180
Oy	782	AAATTTATCAAGACATTTGAGACATAGAAATCTAAATTGGAGATGTGAAGACATACCTTAC	841
Db	181	AAATTTATCAAGACATTTGAGACATAGAAATCTAAATTGGAGATGTGAAGACATACCTTAC	240
Oy	842	ACAGAAATTCCTACTCCCGAAGTGAACAAAGCAAGCACTTACTGTTTACAGTATGATGATC	901
Db	241	AGAGAAATTCCTACTCCCGAAGTGAACAAAGCAAGCACTTACTGTTTACAGTATGATGATC	300
Oy	902	AGAAATATAGTACGGGCGCTTCGAGACAAACACATCAAGATCTGGGATAAAAAACATTTGG	961
Db	301	AGAAATATAGTACGGGCGCTTCGAGACAAACACATCAAGATCTGGGATAAAAAACATTTGG	360
Oy	962	AATGCAAGCGAATTTCTACAGGCCATACAGTTTCACTGCTCTGCTCCAGTATGATGACA	1021
Db	361	AATGCAAGCGAATTTCTACAGGCCATACAGTTTCACTGCTCTGCTCCAGTATGATGACA	420
Oy	1022	GAGTGATCTAAACAGATCATCGGATTCACAGCTGAGATGTGGAGTCT - AAATACAGST	1080
Db	421	GAGTGATCTAAACAGATCATCGGATTCACAGCTGAGATGTGGAGTCTAAATACAGST	480
Oy	1081	GAAATCTTAACACGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTTGGCTTCAATAT	1140
Db	481	GAAATCTTAACACGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTTGGCTTCAATAT	540
Oy	1141	GGCATATGCTGACCGCTGCACCAAGAATCTTCACATTTGATTTGGGATATGGGCTCCCA	1200
Db	541	GGCATATGCTGACCGCTGCCTCCCAAGAATCTTCATTTGCTGTATGGGATATGGCTCCCA	600
Oy	1201	ACTGACATTTACCTCCGGAGGGTGTCTGTC - GCACACGAGCTGTGTCAATGTTGTGA	1259
Db	601	CTTGACATTTACCTCCGGAGGGTGTCTGTCGGACACGAGGCTGTGTCAATGTTGTGA	660
Oy	1260	CTTTGATGACAGT 1273	
Db	661	CTTTGATGACAGT 674	

RESULT 7	BE269297	692 bp	mRNA	linear	EST 13-JUL-2000
LOCUS	BE269297				
DEFINITION	BE269297	601186336F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3544549 5',			
ACCESSION	BE269297	mRNA sequence.			
VERSION	BE269297.1	GI:9142915			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 692)				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

FEATURES	Location/Qualifiers
source	1. .692

```

/db_xref="taxon:9606"
/clone_lib="IMAGE:544549"
/clone_lib="NIH_MGC_8"
/issue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Lymph; Vector: pOTB1; Site: 1; XhoI; Site: 2;
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      202 a      134 c      179 g      177 t
ORIGIN
```

```

Query Match      28.3%; Score 608.8; DB 10; Length 692;
Best Local Similarity 97.7%; Pred. No. 2.8e-171;
Matches 649; Conservative 0; Mismatches 12; Indels 3; Gaps 3;
```

```

QY 602 AATGTCACGAGTGAAGCTGTGATGCTGTGGAAGAGCTTATCGAGAGATGTC 661
DB 1 AATGTCACGAGTGAAGCTGTGATGCTGTGGAAGAGCTTATCGAGAGATGTC 60
QY 662 GGACAGATTTCTGTGGAGAGGCTGGCAGAAAGAGAGATGGGACAGTATTATCA 721
DB 61 GGACAGATTTCTGTGGAGAGGCTGGCAGAAAGAGAGATGGGACAGTATTATCA 120
QY 722 AAAACAACTCTGACGGGAATGCTCTCCCACTCTTTTATAGACACTTATCTTA 781
DB 121 AAAACAACTCTGACGGGAATGCTCTCCCACTCTTTTATAGACACTTATCTTA 180
QY 782 AAATATACAGACATATGAGCAATAGATTAATGAGATGGAAGACATATTAC 841
DB 181 AAATATACAGACATATGAGCAATAGATTAATGAGATGGAAGACATATTAC 240
QY 842 AGAGATTCACCTCCGAGAGTGAACAGCAAGAGATTTACTGTTTACATGATGATC 901
DB 241 AGAGATTCACCTCCGAGAGTGAACAGCAAGAGATTTACTGTTTACATGATGATC 300
QY 902 AGAAATATGTAAGGCGCTTGGAGCAACATCAATGATGGAATTAACACATTGG 961
DB 301 AGAAATATGTAAGGCGCTTGGAGCAACATCAATGATGGAATTAACACATTGG 360
QY 962 AATGCAAGCAATTTCTACAGGCAATACAGTTAGTCTCTGCTCCAGTATGATGA 1021
DB 361 AATGCAAGCAATTTCTACAGGCAATACAGTTAGTCTCTGCTCCAGTATGATGA 420
QY 1022 GAGTATCATTAACAGATCATGATTCAGAGTCAAGTGTGGAGTGAATTAACAGCTG 1081
DB 421 GAGTATCATTAACAGATCATGATTCAGAGTCAAGTGTGGAGTGAATTAACAGCTG 480
QY 1082 AAATGCTAAGACGTTGATTCACATTTGGAAGAGATTTGCACTTGGTTCAATATG 1141
DB 481 AAATGCTAAGACGTTGATTCACATTTGGAAGAGATTTGCACTTGGTTCAATATG 540
QY 1142 GCATGATGAGTGAAGT-GCTCCAAAGATCGTTTCAATGCTGTATGAGATATGCTCCCA 1200
DB 541 GCATGATGAGTGAAGT-GCTCCAAAGATCGTTTCAATGCTGTATGAGATATGCTCCCA 599
QY 1201 ACTGACATTAACCTCCGAGAGTGTGCTGTGAGACACGAGTCTGCTCAATGTTAGAG 1260
DB 600 CTGAAACATTAACCTCCGAGAGTGTGCTGTGAGAC-CGAGACTGTGTCAATGAGTGTGCT 658
QY 1261 TTGG 1264
DB 659 TTGG 662
```

```

RESULT 8
LOCUS      B1822845      828 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603040168F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180993 5',
```

```

ACCESSION      mRNA sequence.
VERSION        B1822845
KEYWORDS       B1822845.1 GI:15934395
SOURCE         EST.
ORGANISM       human.
                Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
                1 (bases 1 to 828)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cga@pds-femail.nih.gov
                Tissue Procurement: Life Technologies, Inc.
                CDNA Library Preparation: Life Technologies, Inc.
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLNL1451 row: h column: 18
                High quality sequence stop: 756.
```

```

FEATURES
    source
        1..828
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="IMAGE:5180993"
            /clone_lib="NIH_MGC_115"
            /lab_host="DH10B"
            /note="Organ: pooled brain, lung, testis; Vector:
                pCMV-Sport6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA
                source anonymous pool of 6 male brains, age range 23-27; 1
                male lung, age 27; and 1 male testis, age 69. Library is
                oligo-dT primed and directionally cloned (EcoRV site is
                destroyed upon cloning). Average insert size 1.8 kb,
                insert size range 1-3 kb. Library is normalized and
                enriched for full-length clones and was constructed by C.
                Gruber (Invitrogen). Research Genetics tracking code
                021. Note: this is a NIH_MGC Library."
```

```

BASE COUNT      241 a      187 c      215 g      185 t
ORIGIN
```

Query Match 27.4%; Score 589.4; DB 13; Length 828;
Best Local Similarity 97.5%; Pred. No. 2.1e-165;
Matches 621; Conservative 0; Mismatches 11; Indels 5; Gaps 2;

```

QY 112 TTATGATTTCTCAGAGAGAGAGACTGTAATTAATGGGAACCCCTAGAGATTAATA 171
DB 188 TTCCAGAAATTCCTCAGAGAGAGAGACTGTAATTAATGGGAACCCCTAGAGATTAATA 247
QY 172 CCAGAGAAATTCCTCAGAGAGAGAGACTGTAATTAATGGGAACCCCTAGAGATTAATA 231
DB 248 CCAGAGAAATTCCTCAGAGAGAGAGACTGTAATTAATGGGAACCCCTAGAGATTAATA 307
QY 232 GAAACAGTATGTTTGAACAAGCACTGTAATGAAGTGAAGATTTGTTGGCCAAACAAA 291
DB 308 GAAACAGTATGTTTGAACAAGCACTGTAATGAAGTGAAGATTTGTTGGCCAAACAAA 367
QY 292 CTTGGCAATGGGACCTTCAATGATTTGTTGCCAAAGCAAGGAACTCTGCAAGAGCTAT 351
DB 368 CTTGGCAATGGGACCTTCAATGATTTGTTGCCAAAGCAAGGAACTCTGCAAGAGCTAT 427
QY 352 GAAAAGAAAAGAACTGTGTCAATACCTTTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 411
DB 428 GAAAAGAAAAGAACTGTGTCAATACCTTTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 487
QY 412 GAATTTTGGGAACATCTATATCCAAATGTCATTAACCAATGGGACATTAAGTCG 471
DB 488 GAATTTTGGGAACATCTATATCCAAATGTCATTAACCAATGGGACATTAAGTCG 547
QY 472 TATCTTAACCATGTTGACAGAGATTTCAATGCTGTGCAAGCTGGGAGTTGGAT 531
```


Db 548 TATCTTAAACCTATGTGCGAGAGATTTCATACCTGCTGCCAGCTCGGGATTGGAT 607
Qy 532 CATATGCTGAGACATTTCTGTCTATACCTGATGCGCAATACATATGTGCTGTAACCT 591
Db 608 CATATGCTGAGACATTTCTGTCTATACCTGATGCGCAATACATATGTGCTGTAACCT 667
Qy 592 GTGTGCAAGGAATGCTGACGAGTGCCTGTGATGCGATGCTGATGCGATGCTGATGCG 651
Db 668 GTGTGCAAGGAATGCTGACGAGTGCCTGTGATGCGATGCTGATGCGATGCTGATGCG 727
Qy 652 AATATGCTGAGAGATTTCTGTGAGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 708
Db 728 AATATGCTGAGAGATTTCTGTGAGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 787
Qy 709 CAGTATTTATTCACAAACCAACCTCTGACGCGAATG 745
Db 788 CAGTATTTATTCACAAACCAACCTCTGACGCGAATG 822

RESULT 9
B0948186 1147 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT.8878641 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6464639
DEFINITION 5', mRNA sequence.
ACCESSION B0948186
VERSION B0948186.1 GI:22363664
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1147)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNL1986 row: a column: 24
High quality sequence stop: 501.
Location/Qualifiers
1. 1147
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6464639"
/clone_id="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 324 a 289 c 284 g 250 t
ORIGIN

Query Match 27.3%; Score 586.6; DB 14; Length 1147;
Best Local Similarity 97.6%; Pred. No. 1.7e-164;
Matches 606; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 112 TTATGAATTCCTCAG 171
Db 150 TTCCAGAAATTCCTCAG 209
Qy 172 CCAG 231
Db 210 CCAG 269
Qy 232 GAAACAGTATGTTTAGCAAGACCTCTATGAAAGTGAAGATGGAATTTGTGGCCCAACAAA 291

Db 270 GAAACAGTATGTTTAGCAAGACCTCTATGAAAGTGAAGATGGAATTTGTGGCCCAACAAA 329
Qy 292 CTGCGCAATGAGACATTTCTGTCTATACCTGATGCGCAATACATATGTGCTGTAACCT 351
Db 330 CTGCGCAATGAGACATTTCTGTCTATACCTGATGCGCAATACATATGTGCTGTAACCT 389
Qy 352 GAAAG 411
Db 390 GAAAG 449
Qy 412 GAATTTGTGGAACATCTTATATCCAAATGTGTATTACCAACATGGGCAATTAAGCTG 471
Db 450 GAATTTGTGGAACATCTTATATCCAAATGTGTATTACCAACATGGGCAATTAAGCTG 509
Qy 472 TATCTTAAACCTATGTTGAGAGAGATTTCAATACCTGCTGCCAGCTCGGGAGTTGGAT 531
Db 510 TATCTTAAACCTATGTTGAGAGAGATTTCAATACCTGCTGCCAGCTCGGGAGTTGGAT 569
Qy 532 CATATGCTGAGACATTTCTGTCTATACCTGATGCGCAATACATATGTGCTGTAACCT 591
Db 570 CATATGCTGAGACATTTCTGTCTATACCTGATGCGCAATACATATGTGCTGTAACCT 629
Qy 592 GTGTGCAAGGAATGCTGACGAGTGCCTGTGATGCGATGCTGT- GGAAGAGAGTTATCGA 650
Db 630 GTGTGCAAGGAATGCTGACGAGTGCCTGTGATGCGATGCTGTGGAAGAGAGTTATCGA 689
Qy 651 GAGATGCTGAGACATTTCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
Db 690 GAGATGCTGAGACATTTCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 749
Qy 711 GTATTTATTCACAAACCAAC 731
Db 750 GTATTTATTCACAAACCAAC 770

RESULT 10
B0829037 743 bp mRNA linear EST 22-MAY-2001
LOCUS 602753364F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4906215 5',
DEFINITION mRNA sequence.
ACCESSION B0829037
VERSION B0829037.1 GI:14176624
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 743)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCMB808 row: o column: 16
High quality sequence stop: 703.
Location/Qualifiers
1. 743
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4906215"
/clone_id="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pORF7; Site:1: EcoRI;
Site:2: XhoI; cDNA made by oligo-dT priming."

Query Match	27.0%;	Score 580.6;	DB 12;	Length 743;
Best Local Similarity	96.4%;	Pred. No. 8.7e-163;		
Matches 637;	Conservative 0;	Mismatches 19;	Indels 5;	Gaps 4;

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1- (bases 1 to 752)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@b-riemail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NIGRI), Shitaki Toshiluyki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LMLM10739 row: p column: 03 High quality sequence stop: 728.
FEATURES	Location/Qualifiers
source	1..752
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:4825970"
	/clone_1ib="NIH_MGC_97"
	/lab_host="DH10B"
	/note="Organ: testis; Vector: pluescriptR (modified pluscript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag) ; Oligo-dt primed using primer 5'-gttttttttttttttttttttVNV-3', size-selected for average insert size 2.2 kb and normalized to R0r 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NIGRI, National Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT	238 a 143 c 194 g 177 t
ORIGIN	
Query Match	26.3% Score 566.4; DB:12; Length 752;
Best Local Similarity	98.8%; Pred. No.1.6e-138;
Matches 581; Conservative	0; Mismatches 6; Indels 1; Gaps 1;
OY	115 ATGAATTCCTCGAGAGAGAAGACACTGTATATATGCGCAACCCTCCAGTAAGATTATACCA 174
Db	161 AAGATTCTCTCAGAGAGAAAGACTGTATATATGCGCAACCCTCCAGTAAGATTATACCA 220
OY	175 GAGAAGATTTCACCTTAGACACACACATACAACAGCTGTGCCAGACTCTGCTTTAAACCAAGAA 234
Db	221 GAGAAGATTTCACCTTAGACACACACATACAACAGCTGTGCCAGACTCTGCTTTAAACCAAGAA 280
OY	235 ACAGATATGTTAGACAGAGACTGTCTTGAAAGCTGGAATGTGGTCCAAAACAATACTT 294
Db	281 ACAGATATGTTAGACAGAGACTGTCTTGAAAGCTGGAATGTGGTCCAAAACAATACTT 340
OY	295 GCCAATGGCACTTCCAGTATGATTTGTGCCCAAGCAACGAAAAAATCTTCAGCAAGCATAGAA 354
Db	341 GCCAATGGCACTTCCAGTATGATTTGTGCCCAAGCAACGAAAAAATCTTCAGCAAGCATAGAA 400
OY	355 AAGGAAAAGGAACGTGTGTCCAATATCTTTAGAGACATGTGTCAAGATGATCAAGTGGAA 414
Db	401 AAGGAAAAGGAACGTGTGTCCAATATCTTTAGAGACATGTGTCAAGATGATCAAGTGGAA 460
OY	415 TTGTGTGAAACATCTTATATCCCAATGTGTATTAACCAATGGGACATATAACTCGAT 474
Db	461 TTGTGTGAAACATCTTATATCCCAATGTGTATTAACCAATGGGACATATAACTCGAT 520
OY	475 CTTTAAACCTATGTGGACAGAGAGATTTCAATAC-TGCTCTGCCAGCTCGGGGATTTGGATCA 533
Db	521 CTTTAAACCTATGTGGACAGAGAGATTTCAATAC-TGCTCTGCCAGCTCGGGGATTTGGATCA 580
OY	534 TATGCTGTGAGAACATTTCTGTCATACCTGGAGGCAAAATCACTATGTGTGTGTAACCTTGT 593
Db	581 TATGCTGTGAGAACATTTCTGTCATACCTGGAGGCAAAATCACTATGTGTGTGTAACCTTGT 640
OY	594 GTGCAAGGAATGTACCGAGTGAACCTGTGATGGCATGTGTGGAAGAAGCTTATTCGAGAG 653

|||||
Db 641 GTCCAGGATGGTACCGAGTGCCTCATGCGATGCTGGAAAGCTTATCGAG 700
QY 654 AATGTCAGAGCAGATTCTGTGAGAGAGCCCTGGCAGAAAGAG 701
|||||
Db 701 AATGTCAGAGCAGATTCTGTGAGAGAGCCCTGGCAGAAAGAG 748
RESULT 12
Bg518761 789 bp mRNA linear EST 02-APR-2001
LOCUS Bg518761
DEFINITION 602578346F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3491843 5',
mRNA sequence.
ACCESSION Bg518761 GI:13513945
VERSION Bg518761.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 789)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1A8536 row: k column: 12
High quality sequence stop: 587.
Location/Qualifiers
1. 789
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3491843"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary."
/lab_host="DH10B"
/note="Organ: Lung; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 257 a 176 c 202 g 154 t
ORIGIN

Query Match 24.8%; Score 534.4; DB 12; Length 789;
Best Local Similarity 86.1%; Pred. No. 7e-149;
Matches 661; Conservative 0; Mismatches 101; Indels 6; Gaps 6;

QY 65 CGATTATGACCCGCGGAGCGTCTGCAAGAGACATCAAGTATGATTCCT 124
|||||
Db 1 CGATTATGACCCGCGGAGCGTCTGCAAGAGAGGCGCTTAGTATGATTCCT 60
|||||
QY 125 CAGAGAGAGAGAGAGTATTAATGCGAAGCCCTTAGAGAGATTAATACAGAGAAAT 184
|||||
Db 61 CAGAGAGAGAGAGAGTATTAATGCGAAGCCCTTAGAGAGATTAATACAGAGAAAT 120
|||||
QY 185 CACTAGACAGACATCAAGAGTGTGCGAGACTGCTTAACCAAGAAAGATGTT 244
|||||
Db 121 CACTAGACAGACATCAAGAGTGTGCGAGCTTTGATTAACCAAGACAGATGTC 180
|||||
QY 245 TAGCAAGCACTGCTATGAGAGTGAATGTTGTGGCCAAACAACTTGCCATGGA 304
|||||
Db 181 TAACAAGCACTGCTATGAGAGTGAATGTTGTGGCCAAACAACTTGCCATGGA 240
|||||
QY 305 CTTCAGTATGATTTGCCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 363

|||||
Db 241 CTTCAGAGATGATTTGTGCGCCAGAGCGGAAACCTCTGAGCAAGCTATGAGAGAAAG 300
QY 364 GAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 422
|||||
Db 301 GAGCTGTGTCTCAATATTTTGGAGAGTGTGAGAGTGTGATCAATGAGAAATTTGTA 360
|||||
QY 423 ACATCTATATCCAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 482
|||||
Db 361 ACACCTTATATCCAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
|||||
QY 483 TATGTTGAGAGAGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 542
|||||
Db 421 TATGTTGAGAGAGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
|||||
QY 543 GAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
|||||
Db 481 GAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
|||||
QY 603 ATGTCACCG-AGTGAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 661
|||||
Db 541 ATGTCACCGAGTGAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 600
|||||
QY 662 GGAC-AGATTTCTGTGGA-GAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGATTT 719
|||||
Db 601 GGACGCGGCTACTCTGTGAGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGATTT 660
|||||
QY 720 CAAAAAACAACCTCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
|||||
Db 661 AAAAACAACAACCTCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 720
|||||
QY 780 TAAATTTATCAAGACATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 827
|||||
Db 721 TAAATTTATCAAGACATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 767
|||||

RESULT 13
B1103092 637 bp mRNA linear EST 26-JUN-2001
LOCUS B1103092
DEFINITION 602889321P1 NCI_CGAP_kid14 Mus musculus cDNA clone IMAGE:5044684
5', mRNA sequence.
ACCESSION B1103092
VERSION B1103092.1 GI:14553985
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 637)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1A11122 row: 1 column: 05
High quality sequence stop: 629.
Location/Qualifiers
1. 637
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5044684"
/clone_lib="NCI_CGAP_kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT."

Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI-CCAB Library. 1"

ACCESSION B145545
VERSION B145545.1 GI:15246201
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 888)
Mammalia; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1639 row: d column: 19
High quality sequence start: 5
High quality sequence stop: 684.

FEATURES
source location/Qualifiers
1..888

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5253090"
/tissue_type="NCL CGAP Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; Sail; Site: 2; Not: Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lohar Hennighausen/Robin Humphreys, NIH"

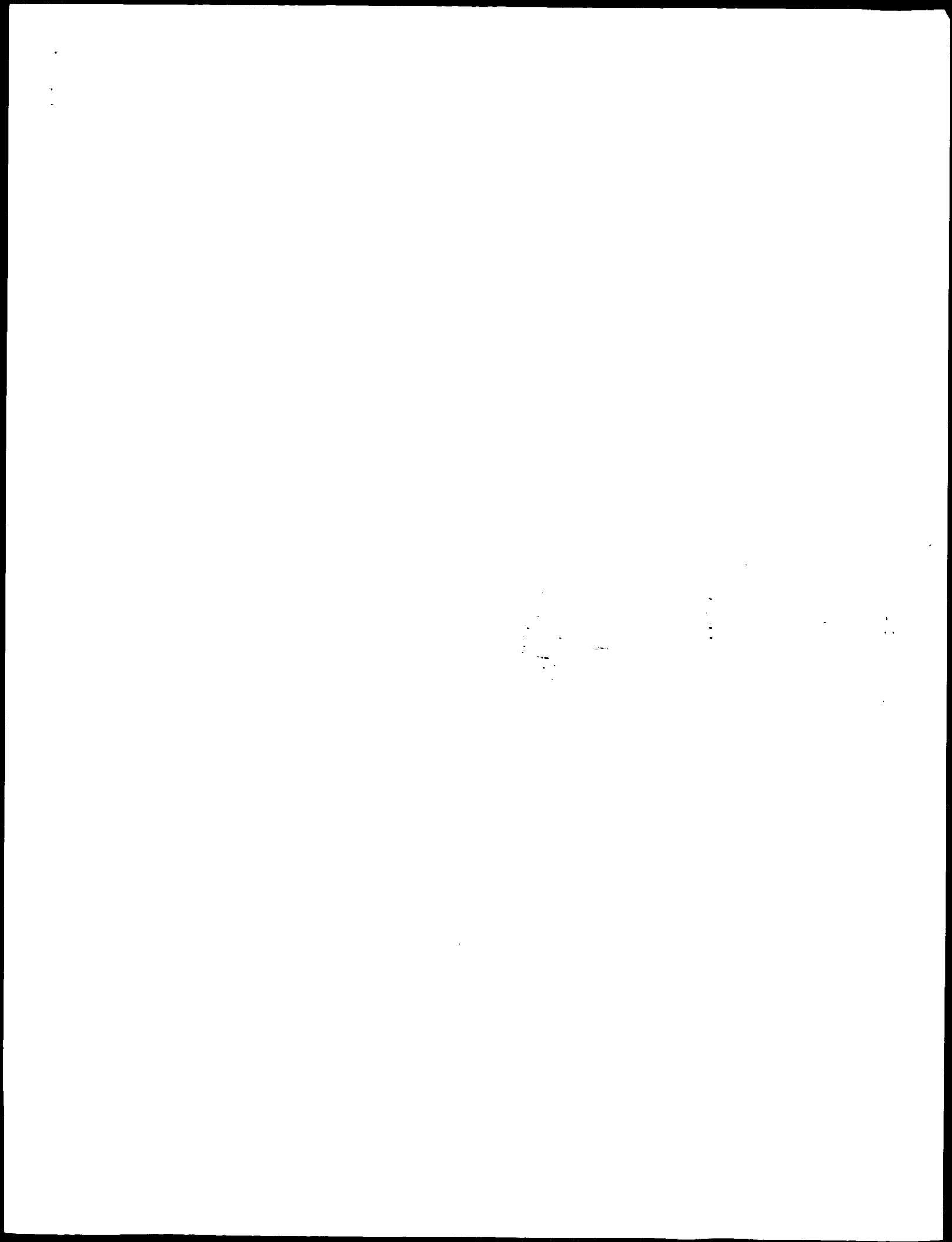
BASE COUNT 219 a 218 c 248 g 203 t
ORIGIN

Query Match 24.6%; Score 528.4; DB 13; Length 888;
Best Local Similarity 81.6%; Pred. No. 4.8e-147;
Matches 725; Conservative 0; Mismatches 141; Indels 22; Gaps 9;

QY 854 GCCAAGTGAACAGCAAGAGGATTTAGCTTACAGTATGATGATCAGAAATAGTAA 913
DB 1 GCCGAGTGAACAGCAAGAGGATTTAGCTTACAGTATGATGATCAGAAATAGTAA 60
QY 914 GCGGCTTCGAGACACACATCAAGATCTGGATTAACACATTTGGAATGCAAGCGAA 973
DB 61 GCGGCTTCGAGACACACATCAAGATCTGGATTAACACATTTGGAATGCAAGCGAA 120
QY 974 TTCTCAGAGGCGATACAGTTCCTCTCTCTCTCCAGTATGATGATGATGATGAT 1033
DB 121 TTCTCAGAGGCGATACAGTTCCTCTCTCTCTCCAGTATGATGATGATGATGAT 180
QY 1034 CAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093
DB 181 CAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 1094 CGTTGATTCACCATTTGGA- GCAATTCGACATTCGCTTCAATTAATGGCATGATG 1152
DB 241 CATTGATTCACCATTTGGA- GCAATTCGACATTCGCTTCAATTAATGGCATGATG 300
QY 1153 ACCTGCTCAAGATGCTTCATGCTGTATGGATATGGCTCCCAATGACATTAAC 1212
DB 301 ACCTGCTCAAGATGCTTCATGCTGTATGGATATGGCTCCCAATGACATTAAC 360
QY 1213 CTCGAGAGGTGCTGTGAGACACGAGCTCTGCAATGTTGAGACTTGTATGACAG 1272
DB 361 CTCGAGAGGTGCTGTGAGACACGAGCTCTGCAATGTTGAGACTTGTATGACAG 420

QY 1273 TACATGTTTCTGCATCTGGGATAGAACTATPAAAGTATGAAACACAGTACTTTGAA 1332
DB 421 TACATGTTTCTGCATCTGGGATAGAACTATPAAAGTATGAAACACAGTACTTTGAA 480
QY 1333 TTTGTAAGACCTTAAATGAGACACAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1392
DB 481 TTTGTAAGACCTTAAATGAGACACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 1393 CTGTAAGTGTGAGTGTCTCATCTGACAACTATCAGATATGAGGAGGAGGAGGAGG 1451
DB 541 CTGTAAGTGTGAGTGTCTCATCTGACAACTATCAGATATGAGGAGGAGGAGGAGGAGG 600
QY 1452 ATGTTTACGAGTGTAGAAAGGACATGAGA-ATTGCTGCTTGTATTTGATTAAC- 1509
DB 601 ATGCTTGGAGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 1510 AAGAGATATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1564
DB 661 AAGAGATATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 1565 CTGCTTTGGACCCCGGCTGCTCTGCGAGGAGACACTCTGTCTAC---GGACCTTGTGAG 1620
DB 721 CTGCTTTGGACCCCGGCTGCTCTGCGAGGAGACACTCTGTCTAC---GGACCTTGTGAG 780
QY 1621 C---ATTCCGGAAGACTTTTGTGACTAC---AGTTGATGATTCAG-ATTGTCACT 1671
DB 781 CCAATTTCTGGCAAAAAGTTTCCGACCTCCAGTTTGGATGAGCTTCCAGCATGCTGACG 840
QY 1672 AGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1719
DB 841 AGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888

Search completed: February 22, 2003, 00:36:50
Job time : 2816 secs



|||||
Db 241 TGTTCAGAACCTGCTATGAGACAGTGAATGTTGTGGCCAAACAACTTGCCAT 300
Qy 301 GGCACTTCAGTATGATGTCGCCAAGCAAGGAACTCAGCAACCTATGAAAAGAA 360
Db 301 GGCACTTCAGTATGATGTCGCCAAGCAAGGAACTCAGCAACCTATGAAAAGAA 360
Qy 361 AAGGAACTGTGTGTCAAATACCTTGTGACAGTGTGTGACAGTCAATCAATTTGTG 420
Db 361 AAGGAACTGTGTGTCAAATACCTTGTGACAGTGTGTGACAGTCAATCAATTTGTG 420
Qy 421 GAACATCTTATATCCCAATATGTGATACCAACATGGCAATTAATCTGTATTTAA 480
Db 421 GAACATCTTATATCCCAATATGTGATACCAACATGGCAATTAATCTGTATTTAA 480
Qy 481 CCTATGTTGAGAGAGATTCATTAACCTGCTGTGACCTGGGGATTTGATCAATGCT 540
Db 481 CCTATGTTGAGAGAGATTCATTAACCTGCTGTGACCTGGGGATTTGATCAATGCT 540
Qy 541 GAGAACATTTCTGATACCTGATGCTGCAATCACTATGCTGCTGAACTTGTGCAAG 600
Db 541 GAGAACATTTCTGATACCTGATGCTGCAATCACTATGCTGCTGAACTTGTGCAAG 600
Qy 601 GAATGTGACGAGTACCTGCTGATGCTGCTGTGGAAGACTTATCGAGAGATGCTC 660
Db 601 GAATGTGACGAGTACCTGCTGATGCTGCTGTGGAAGACTTATCGAGAGATGCTC 660
Qy 661 AGGACAGATTTCTGTCGAGAGAGCTGGCAGAGCAAGAGATGGGAGACATTTATTC 720
Db 661 AGGACAGATTTCTGTCGAGAGAGCTGGCAGAGCAAGAGATGGGAGACATTTATTC 720
Qy 721 AAAAACAACTCTCTGACGGGAAATGCTCTCCCACTTTTATAGACACTTTATCC 780
Db 721 AAAAACAACTCTCTGACGGGAAATGCTCTCCCACTTTTATAGACACTTTATCC 780
Qy 781 AAAATTAACAAGCAATTTGAGCAATTAATGATGATGATGATGATGATGATGAT 840
Db 781 AAAATTAACAAGCAATTTGAGCAATTAATGATGATGATGATGATGATGATGAT 840
Qy 841 CAGAGATTCCTGCTGCGAAGTGAACAGCAAGAGATTACTGTTTACAGTATGAT 900
Db 841 CAGAGATTCCTGCTGCGAAGTGAACAGCAAGAGATTACTGTTTACAGTATGAT 900
Qy 901 CAGAAAAATGTAAGGCGCTTGAGACAAACACATCAAGATCTGGATAAAAACATTTG 960
Db 901 CAGAAAAATGTAAGGCGCTTGAGACAAACACATCAAGATCTGGATAAAAACATTTG 960
Qy 961 GAATGCAAGGAAATTCACAGGCGCATACAGTTCAGTCTCTGTCTCCAGTATGATG 1020
Db 961 GAATGCAAGGAAATTCACAGGCGCATACAGTTCAGTCTCTGTCTCCAGTATGATG 1020
Qy 1021 AGAGTATCATTAACAGATCATCGATTCACAGGTCAGAGTGGGATTAATACAGGT 1080
Db 1021 AGAGTATCATTAACAGATCATCGATTCACAGGTCAGAGTGGGATTAATACAGGT 1080
Qy 1081 GAAATCTTAAACAGTTGATTCACATTTGTAAGCAGTTCTGCACTTGCTTCAATAT 1140
Db 1081 GAAATCTTAAACAGTTGATTCACATTTGTAAGCAGTTCTGCACTTGCTTCAATAT 1140
Qy 1141 GGCATATGTCAGCTGCTGCAAAAGTCTTCATTTGATGATGATGATGATGATGAT 1200
Db 1141 GGCATATGTCAGCTGCTGCAAAAGTCTTCATTTGATGATGATGATGATGATGAT 1200
Qy 1201 ACTGACATTAACCTCGAGGCTGTGTCGAGACACCGAGCTGCTGCAATGTTGATAG 1260
Db 1201 ACTGACATTAACCTCGAGGCTGTGTCGAGACACCGAGCTGCTGCAATGTTGATAG 1260
Qy 1261 TTTGATGACAGTATGATTTTCTGATCTGAGGATAGAACTATAAAGTATGAAACACA 1320
Db 1261 TTTGATGACAGTATGATTTTCTGATCTGAGGATAGAACTATAAAGTATGAAACACA 1320
Qy 1321 AGTACTTGTGATTTGTAGAGACCTTAATGACCAAAACGAGCTTGCCTGTTGCAG 1380
|||||

Db 1321 AGTACTTGTGATTTGTAGAGACCTTAATGACCAAAACGAGCTTGCCTGTTGCAG 1380
Qy 1381 TACAGGACAGCTGTGTAGTGTGATCTGATCAACACATATCAGATTAAGGACATA 1440
Db 1381 TACAGGACAGCTGTGTAGTGTGATCTGATCAACACATATCAGATTAAGGACATA 1440
Qy 1441 GAATGTGTGATTTTACAGTGTAGAGGCGCATGAGGAATTTGTGCTGTTATTTGA 1500
Db 1441 GAATGTGTGATTTTACAGTGTAGAGGCGCATGAGGAATTTGTGCTGTTATTTGA 1500
Qy 1501 TTTGATTAACAGAGATATGTCAGTGGGCTTATGATGAAAAATTAAGTGTGGATCT 1560
Db 1501 TTTGATTAACAGAGATATGTCAGTGGGCTTATGATGAAAAATTAAGTGTGGATCT 1560
Qy 1561 GTGGCTGCTTTGGACCCCGCTCTCTGAGGAGACATCTGTCTACGAGACCTTTGTGAG 1620
Db 1561 GTGGCTGCTTTGGACCCCGCTCTCTGAGGAGACATCTGTCTACGAGACCTTTGTGAG 1620
Qy 1621 CATTCGGAAGAGTTTTCGATACAGTTTGTGATGAAATTCAGATTTGTCATGATTCAT 1680
Db 1621 CATTCGGAAGAGTTTTCGATACAGTTTGTGATGAAATTCAGATTTGTCATGATTCAT 1680
Qy 1681 GATGACAAATCTGATCTGAGACTTCTTAATGATGATGATGATGATGATGATGAT 1740
Db 1681 GATGACAAATCTGATCTGAGACTTCTTAATGATGATGATGATGATGATGATGAT 1740
Qy 1741 CGTTCCCTCTCTGAGACATACACCTACATCTCCAGATTAATTAACATACATGAT 1800
Db 1741 CGTTCCCTCTCTGAGACATACACCTACATCTCCAGATTAATTAACATACATGAT 1800
Qy 1801 ACTTCCGAGAGACCATTAAGTGGGATTTTAACGATGTCAGATTAACAGATTAAGC 1860
Db 1801 ACTTCCGAGAGACCATTAAGTGGGATTTTAACGATGTCAGATTAACAGATTAAGC 1860
Qy 1861 AACAACAGTAACATCAACATCACTGCGCAGTTTCCCTGAGTACGAGAGAGAGGCTT 1920
Db 1861 AACAACAGTAACATCAACATCACTGCGCAGTTTCCCTGAGTACGAGAGAGAGGCTT 1920
Qy 1921 TGAGACTCTGTTGGGACAGAGTGTGTCAGTGTGCGCCGAGAGGCTTACTACAGACA 1980
Db 1921 TGAGACTCTGTTGGGACAGAGTGTGTCAGTGTGCGCCGAGAGGCTTACTACAGACA 1980
Qy 1981 ACTGACTCTTCAGTGTGCTGATCAGAGATGCTTATCAATTTGATGATGATGATGAT 2040
Db 1981 ACTGACTCTTCAGTGTGCTGATCAGAGATGCTTATCAATTTGATGATGATGATGAT 2040
Qy 2041 TTTTAACTCCCT 2100
Db 2041 TTTTAACTCCCT 2100
Qy 2101 CAGCAAAAGTGACTTAAATATATTAGTGTGTTCCAGAAAAA 2151
Db 2101 CAGCAAAAGTGACTTAAATATATTAGTGTGTTCCAGAAAAA 2151

RESULT 2
US-10-023-530-1
: Sequence 1, Application US/10023530
: Publication No. US2003007956A1
: GENERAL INFORMATION:
: APPLICANT: LAGRAN, Pierre
: APPLICANT: BERNARD, Richard
: APPLICANT: BIOT, Guillaume
: APPLICANT: LASSOT, Irina
: TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP
: FILE REFERENCE: B4717A
: CURRENT APPLICATION NUMBER: US/10/023, 530
: CURRENT FILING DATE: 2002-04-22
: PRIOR APPLICATION NUMBER: 60/256, 276
: PRIOR FILING DATE: 2000-12-18
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1

[illegible]

RESULT 4

Sequence: 09-864-761-2778/c
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Gene Expression Analysis by Microarray

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2778

LENGTH: 479

TYPE: DNA

ORGANISM: Homo sapiens

FEATURES:

OTHER INFORMATION: MAP TO AP000252.1.

OTHER INFORMATION: EXPRESSED IN HBLLI00, SIGNAL = 3.4

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9

```

: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
US-09-864-761-2778

Query Match      12.6%; Score 272; DB 10; Length 479;
Best Local Similarity 74.2%; Pred. No. 3,36-73;
Matches 359; Conservative 0; Mismatches 120; Indels 5; Gaps 1;

OY 1146 CATGCTGACCTGCTCCAAAGATCGTTCATGCTGCTGATGAGGATATGCGCTCCCACTGA 1205
DB 479 GATGGTGAACCTGTTCCAAAGACCGCTGCATGCTGCTGAGACATGCGTTCTGGACGGA 420
OY 1206 CATTCACCTCCGAGAGGGTGGTGGACACCGACGCTGCTGCATGTTGTAAGCTTTGA 1265
DB 419 TATCACTTAGTGGTGTGTCCTGTTGCTGCTACTGACGCTGCTGTCAACGTAAGACTTTGA 360
OY 1266 TGACAGTAGCATTTGTTCTGCATCTGGGATAGACTATAAAGGTATGGAACACAGTAC 1325
DB 359 CGAGAAATACATCGTGTCTGCTCTGCTGTGACAGGCCATCAAACTGAGACATGAGAC 300
OY 1326 TTGCTAATTTGTAAGAGCCTAAATGAGACAAACGAGACGCTTCCCTGTTGACGTACAG 1385
DB 299 CTGTGAATTTGCTGTACTCTCAATGGGCAAGGAG-----TCTGTCTCCAGTACAG 245
OY 1386 GGACAGCGCTGGTAGTGAATGGCTCATCTGACACACTATCAGATTATGAGACATAGAAATG 1445
DB 244 GGATGGCGCTGGTGTGTATGATGATCATCATCATACATTAAGCTCTTGATATTGAATG 185
OY 1446 TGGTGCATGTTTACGAGTGTAGAAAGCCATGAGGAATGGTGGCTGTATTCATTTGA 1505
DB 184 TGGCTCCCTGTTTAAAGTCCATGAGGGACATGAAGAAATGGTCCGATGCATCCAGTTTGA 125
OY 1506 TAACAAAGAGATAGTACATGGGCTATGATGAGAAAATTTAAAGTGGAGCTTTGGGC 1565
DB 124 TAACAAGAGATGCTGACAGCCGGGCTATGATGGGAAAATTTAAATTATGAGACTTGCAAGC 65
OY 1566 TGCYTTGAGCCCCGCTGCTCTGTCAGAGGACACTGTCTACGAGCCCTTGTGAGCAATTC 1625
DB 64 CGCTCTTGACCCCTGAGAGCCCGACAGACATTTGTTGGCACATTTGGTGAACATTC 5
OY 1626 CGGA 1629
DB 4 TEGA 1

RESULT 5
US-09-864-761-19492/c
: Sequence 19492, Application US/09864761
: Patient No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aemica-X-1
: CURRENT APPLICATION NUMBER: US/09/864, 761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667

```

```

1 PRIOR FILING DATE: 2001-01-30
2 PRIOR APPLICATION NUMBER: PCT/US01/00664
3 PRIOR FILING DATE: 2001-01-30
4 PRIOR APPLICATION NUMBER: PCT/US01/00669
5 PRIOR FILING DATE: 2001-01-30
6 PRIOR APPLICATION NUMBER: PCT/US01/00665
7 PRIOR FILING DATE: 2001-01-30
8 PRIOR APPLICATION NUMBER: PCT/US01/00668
9 PRIOR FILING DATE: 2001-01-30
10 PRIOR APPLICATION NUMBER: PCT/US01/00663
11 PRIOR FILING DATE: 2001-01-30
12 PRIOR APPLICATION NUMBER: PCT/US01/00662
13 PRIOR FILING DATE: 2001-01-30
14 PRIOR APPLICATION NUMBER: PCT/US01/00661
15 PRIOR FILING DATE: 2001-01-30
16 PRIOR APPLICATION NUMBER: PCT/US01/00670
17 PRIOR FILING DATE: 2001-01-30
18 PRIOR APPLICATION NUMBER: US 60/234,687
19 PRIOR FILING DATE: 2000-09-21
20 PRIOR APPLICATION NUMBER: US 09/608,408
21 PRIOR FILING DATE: 2000-06-30
22 PRIOR APPLICATION NUMBER: US 09/774,203
23 PRIOR FILING DATE: 2001-01-29
24 NUMBER OF SEQ ID NOS: 49117
25 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
26 SEQ ID NO 19492
27 LENGTH: 370
28 TYPE: DNA
29 ORGANISM: Homo sapiens
30 FEATURE:
31 OTHER INFORMATION: MAP TO AP000252.1
32 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.4
33 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
34 OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.1
35 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8
36 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
37 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
38 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
39 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3
40 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
41 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
42 OTHER INFORMATION: NT HIT: AB014596.1, EVALUE 0.006+00
43 OTHER INFORMATION: EST_HUMAN HIT: AL040743.1, EVALUE 1.00e-102
44 OTHER INFORMATION: SWISSPROT HIT: Q91854, EVALUE 3.00e-46
45 JS-09-864 -761.19492

```

[illegible]

```

OY      1387  GACAGGCTGGTAGT 1400
      || ||||| ||
Db      14  GATGGCTGTTGT 1

RESULT 6
US-09-960-352-2934
; Sequence 2934, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nenping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2934
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (261)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 13-LIB3058-049-Q1-K1-D1
US-09-960-352-2934

```

```

Query Match similarity      8.4%; Score 180.2; DB 10; Length 319;
Best Local Similarity     93.1%; Pred. No. 3,9e+45;
Matches 188; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY    1053 GCCTCAGAGTGTGGATGTAATAATACAGGTGAATGCTAAACACGTTGATTACCATTGTGCA 1112
       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB    118 GTTATGGGTGTGGATGTAATAATACAGGTGAATGCTGAACACATTGATTACCATTTGGA 177
OY    1113 AGCAATTTGCACACTTGCGTTTCATTAATGGCATGATGTGGAGCTCTGCCAAGAATCGTTC 1172
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB    178 AGCAATTTGCACCTTGCGTTTCATTAATGGCATGATGTGGAGCTCTGCCAAGAATCGTTC 237
OY    1173 CATTCCTGATGGATATGAGCCCTCCCAACTGACATTAACCCCTCGAGGGTCTGTGCGG 1232
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB    238 CATTGACGATBTGGATATGAGCCCTNCCCAACTGACATCACCTTCGACAGSGTCTGTGCG 297
OY    1233 ACACCGAGCTGCTGTCAATGTT 1254
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB    298 ACACCGAGCTGCTCAATGTT 319

RESULT 7
US-09-213-888-20
/ Sequence 20, Application US/09213888A
/ Patent No. US20020164683A1
/ GENERAL INFORMATION:
/ APPLICANT: Gurney, Mark E.
/ APPLICANT: Ili, Jinhe
/ APPLICANT: Pauley, Adele M.
/ APPLICANT: Pharmacia & Upjohn Company
/ TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
/ FILE REFERENCE: Encode Them
/ CURRENT APPLICATION NUMBER: US/09/213,888A
/ CURRENT FILING DATE: 1998-12-17
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 20
/ LENGTH: 1881
/ TYPE: DNA
/ ORGANISM: Artificial Sequence

```

```

RESULT 8
US-09-328-877A-20
; Sequence 20, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27

```

RESULT 9
US-09-213-888-26
Sequence 26, Application US/09213888A
Patent No. US20020164683A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Li, Jinhe
APPLICANT: Pauley, Adele M.
TITLE OF INVENTION: Human Sel-10 Polypeptides and polynucleotides that

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:55:27 ; Search time 17 Seconds
(without alignments)
984.803 Million cell updates/sec

Title: US-09-601-168b-2
Perfect score: 3034
Sequence: 1 MDPAAVLQEKALKEFMSSE.....PAAQAEPPSPSRITYTISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2582.5	85.1	517	1	US-08-190-802A-30
2	2582.5	85.1	517	4	US-08-477-346-30
3	2582.5	85.1	517	4	US-08-473-089-30
4	2582.5	85.1	517	4	US-08-487-072A-30
5	545	18.0	640	4	US-09-177-165A-30
6	520	17.1	587	3	US-08-899-578-2
7	399	13.2	779	4	US-08-190-802A-32
8	399	13.2	779	4	US-08-477-346-32
9	399	13.2	779	4	US-08-473-089-32
10	399	13.2	779	4	US-08-487-072A-32
11	399	13.2	779	4	US-09-177-165A-29
12	392	12.9	732	4	US-08-914-999-8
13	354	11.7	409	2	US-08-283-917-3
14	354	11.7	409	2	US-08-961-716-3
15	354	11.7	410	2	US-08-283-917-9
16	354	11.7	410	2	US-08-961-716-9
17	339.5	11.2	409	1	US-08-190-802A-51
18	339.5	11.2	409	4	US-08-477-346-51
19	339.5	11.2	409	4	US-08-473-089-51
20	339.5	11.2	409	4	US-08-487-072A-51
21	337	11.1	1146	4	US-08-914-999-6
22	321.5	10.6	514	1	US-08-190-802A-66
23	321.5	10.6	514	4	US-08-477-346-66
24	321.5	10.6	514	4	US-08-473-089-66
25	321.5	10.6	514	4	US-08-487-072A-66
26	318	10.5	422	1	US-08-190-802A-52
27	318	10.5	422	4	US-08-477-346-52

28	318	10.5	422	4	US-08-473-089-52	Sequence 52, Appl
29	318	10.5	422	4	US-08-487-072A-52	Sequence 52, Appl
30	313.5	10.3	209	3	US-08-899-578-6	Sequence 6, Appl
31	306	10.1	704	1	US-08-188-582-5	Sequence 5, Appl
32	306	10.1	704	1	US-08-646-715-5	Sequence 5, Appl
33	305	10.1	704	1	US-08-190-802A-62	Sequence 62, Appl
34	305	10.1	704	2	US-08-308-818-3	Sequence 3, Appl
35	305	10.1	704	4	US-08-477-346-62	Sequence 62, Appl
36	305	10.1	704	4	US-08-473-089-62	Sequence 62, Appl
37	305	10.1	704	4	US-08-487-072A-62	Sequence 62, Appl
38	299.5	9.9	1194	4	US-09-092-508-2	Sequence 2, Appl
39	299.5	9.9	1194	4	US-09-435-115-2	Sequence 2, Appl
40	299.5	9.9	1194	4	US-09-089-023-26	Sequence 26, Appl
41	299.5	9.9	1194	4	US-09-098-310-2	Sequence 2, Appl
42	299.5	9.9	1205	4	US-09-092-508-16	Sequence 16, Appl
43	299.5	9.9	1205	4	US-09-435-115-16	Sequence 16, Appl
44	298.5	9.8	704	1	US-08-188-582-18	Sequence 18, Appl
45	298.5	9.8	704	1	US-08-646-715-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-190-802A-30
Sequence 30, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0860
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO.: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BETA TRCP, Flg. 13
US-08-190-802A-30
Query Match 85.1%; Score 2582.5; DB 1; Length 517;
Best Local Similarity 91.4%; Pred. No. 2.9e-266;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;
18 SSREDCNNGEPKRIIPKNSLRQIYNSCARCLNQEIVCLASTAMTEKCVAKTKLAN 77

PRINTED

Db 13 ASEREDCNDEPPRKITEKNTLRO-----TKLAN 42
Qy 78 GTSSMIVPRQKRLSASYEKEKELCVKYEQWSESDQVEFEVHLISQCHYOHGHTNSYLK 137
Db 43 GTSSMIVPRQKRLSANYEKEKELCVKYEQWSESDQVEFEVHLISQCHYOHGHTNSYLK 102
Qy 138 PMLQRFETALPARGLDHAENILSYLDAKSLCAELVCKEYRYVTSQGLMKKLIERNV 197
Db 103 PMLQRFETALPARGLDHAENILSYLDAKSLCAELVCKEYRYVTSQGLMKKLIERNV 162
Qy 198 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 257
Db 163 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 222
Qy 258 QRHCSEFSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLMTGHTSVCLQYDE 317
Db 223 QRHCSEFSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLMTGHTSVCLQYDE 282
Qy 318 RVITGSSDSTVRVMDVNTGEMNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 377
Db 283 RVITGSSDSTVRVMDVNTGEMNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 341
Qy 378 TDITLRVLYVGHRAAVNVVDFDDKYIVSASGDRITKVMNTSTCEVFTLNGHKGRIACLO 437
Db 342 TDITLRVLYVGHRAAVNVVDFDDKYIVSASGDRITKVMNTSTCEVFTLNGHKGRIACLO 401
Qy 438 YRDLIVVSSGSDNTIRLMDIECGACLRVLEGEHELVRCIRFDNKRIVSGAYDGKIKYWL 497
Db 402 YRDLIVVSSGSDNTIRLMDIECGACLRVLEGEHELVRCIRFDNKRIVSGAYDGKIKYWL 461
Qy 498 VALDPAPAGTLCRLTVLHSGRVFRLQDFEFOIVSSSHDPTILIMDFLNDP 550
Db 462 VALDPAPAGTLCRLTVLHSGRVFRLQDFEFOIVSSSHDPTILIMDFLNDP 514

RESULT 2
US-08-477-346-30
Sequence 30, Application US/08477346
Patent No. 6262023
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 30:

2-6-95
1-10-95

SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-477-346-30

Query Match 85.1%; Score 2582.5; DB 4; Length 517;
Best Local Similarity 91.4%; Pred. No. 2.9e-266;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

Db 13 ASEREDCNDEPPRKITEKNTLRO-----TKLAN 42
Qy 78 GTSSMIVPRQKRLSASYEKEKELCVKYEQWSESDQVEFEVHLISQCHYOHGHTNSYLK 137
Db 43 GTSSMIVPRQKRLSANYEKEKELCVKYEQWSESDQVEFEVHLISQCHYOHGHTNSYLK 102
Qy 138 PMLQRFETALPARGLDHAENILSYLDAKSLCAELVCKEYRYVTSQGLMKKLIERNV 197
Db 103 PMLQRFETALPARGLDHAENILSYLDAKSLCAELVCKEYRYVTSQGLMKKLIERNV 162
Qy 198 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 257
Db 163 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 222
Qy 258 QRHCSEFSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLMTGHTSVCLQYDE 317
Db 223 QRHCSEFSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLMTGHTSVCLQYDE 282
Qy 318 RVITGSSDSTVRVMDVNTGEMNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 377
Db 283 RVITGSSDSTVRVMDVNTGEMNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 341
Qy 378 TDITLRVLYVGHRAAVNVVDFDDKYIVSASGDRITKVMNTSTCEVFTLNGHKGRIACLO 437
Db 342 TDITLRVLYVGHRAAVNVVDFDDKYIVSASGDRITKVMNTSTCEVFTLNGHKGRIACLO 401
Qy 438 YRDLIVVSSGSDNTIRLMDIECGACLRVLEGEHELVRCIRFDNKRIVSGAYDGKIKYWL 497
Db 402 YRDLIVVSSGSDNTIRLMDIECGACLRVLEGEHELVRCIRFDNKRIVSGAYDGKIKYWL 461
Qy 498 VALDPAPAGTLCRLTVLHSGRVFRLQDFEFOIVSSSHDPTILIMDFLNDP 550
Db 462 VALDPAPAGTLCRLTVLHSGRVFRLQDFEFOIVSSSHDPTILIMDFLNDP 514

RESULT 3
US-08-473-089-30
Sequence 30, Application US/08473089
Patent No. 6342368
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

2-6-95
1-10-95

```

1 SOFTWARE: Patentin Release #1.0, Version #1.25
2
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/08/473,089
5 FILING DATE: 07-JUN-1995
6 CLASSIFICATION: 435
7
8 ATTORNEY/AGENT INFORMATION:
9 NAME: MURASHIGE, KATE H.
10 REGISTRATION NUMBER: 29,959
11 REFERENCE/DOCKET NUMBER: 2550-0025.22
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (202) 887-1500
14 TELEFAX: (202) 887-0763
15
16 INFORMATION FOR SEQ ID NO: 30:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 517 amino acids
19 TYPE: amino acid
20 TOPOLOGY: unknown
21 MOLECULE TYPE: peptide
22 HYPOTHETICAL: NO
23 ANTI-SENSE: NO
24 ORIGINAL SOURCE:
25 INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
26
27 US-08-473-089-30

```

QY 498 VAALDPRAPACTLCTRLTVEHSGVFRLOPDEFOIVSSSHDDTLTMDPLNDP 550
Db 462 VAALDPRAPACTLCTRLTVEHSGVFRLOPDEFOIVSSSHDDTLTMDPLNDP 514

RESULT 5
US-09-177-165A-30
; Sequence 30, Application US/09177165A
; Patent No. 6426205
; GENERAL INFORMATION:
; APPLICANT: Tyers, Mike
; APPLICANT: Williams, Andrew
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
; FILE REFERENCE: 11757.10US01
; CURRENT APPLICATION NUMBER: US/09/177,165A
; CURRENT FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/092,443
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/063,254
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 640
; TYPE: PRF
; ORGANISM: Saccharomyces cerevisiae
US-09-177-165A-30

Query Match 18.0%; Score 545; DB 4; Length 640;
Best Local Similarity 28.9%; Pred. No. 6e-49;
Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

QY 34 IPEKNSLRQTYNSCARCLN-QETVCLASTAMKTEKCVAK-TKLANGSTMTYPRKGRKLS 91
Db 86 LPEKNTKPCYRNPDPLOPPTTACCKOLKRTQELNANAKLPLOBOBDIHIIISKTS 145
QY 92 ASIEKEKELCVKYPEQMSSEDOVEFEVHLISOMCHYOHCHINSYKLPMLORDFTALPAR 151
Db 146 NSNDKIRKL-----ILDGILSTSCFPOLSYISLVTYHMKIDPISLTP-- 188
QY 152 GLDHIAENILSYDAKSICAEELVCKEYKRVYSDGMILMKLLIRMYRTTSLMRGLAERR- 210
Db 189 --GELSKILSYLDCOSLCAKTRCKWOLADDDVWYHMCROH-----DRKC 236
QY 211 --GMOYLEFKNK-----PPDGNAPPNSFYRALYPRKIIDITETISNMRGHRSLQ 258
Db 237 PNCGWGLPILHMKRARIQONSTGSSSNADIQOTTPRWKVIYERKVESNMNRKG----- 291
QY 259 RIRGRSEKSK-----GYCLOYDDOKIVSGIRNDTIKIMDKNTLECKRIILTGHTSVYLCQ 314
Db 292 --HCRIOEFKHMVGVLLOFNRLFTGSDYSTIGIMDLFTGKILRRLSHSDGVKTYL 349
QY 315 YDRVILITGSSDSTVVRWMDNGEMNLTLIHCEAVLHLRFNNGMAYVTSKRSIAVWVM 374
Db 350 FDKRKILITSLDKTIRVWNTITGECISTYRGHSDVSLVSTOKVIVSGSAUKTYKAVWHV 409
QY 375 ASFTDITLRLVGVHRAAVNVYDFEDDKYL--VSASGDRITIKWNTSTCFVFTLNGH-- 429
Db 410 ESRTCYTLR-----GHEWYNVCYKILHPSKSPSCSCSDDTLIRMDIRTNCLAVFRGHVQ 465
QY 430 KRGIACLOYRD--RLVSSGSS----- 448
Db 466 VQKILPTLTKIDENLATDNTSDGSSFPDDPTMTDGADESDEPSNDETVLIDENIPYTHL 525
QY 449 -----DNTRLMDIEGACILRYLEGEHELVRICIRFNDKRIYSGAVDGKIKVMDIYAALDP 503
Db 526 LSGGLDNTIKIMDVYKTKGKIRTFQFGHVEGVMADIADNFIISSGSHDGIKAVMDLQSG--- 582
QY 504 KAPAGTLCRLTVEHSGVFRLOPDEFOIVSSSHDDTL 541
Db 583 -----KCHHTF---NGR---RLQRETOHTOTQSLGDKV 609

RESULT 6
US-08-899-578-2
; Sequence 2, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JPM/AMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-899-578-2

Query Match 17.1%; Score 520; DB 3; Length 587;
Best Local Similarity 28.8%; Pred. No. 2.4e-46;
Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

QY 72 KTKLANGSSMT-----VPKQK--LSASYKEKEL-----CYKYEQMSSEDOVE 115
Db 35 ESSYSNGSSSSSYNADKLSSSRPLQHKIDLSASPSRRNDLNPVEHLIALFQDLSSARQMD 94
QY 116 FVEHLISOMCHYOHCHINSYKLPMLORDFTALPARGLDHIAENILSYDAKSICAEELV 175
Db 95 AFTRLQSSNMNTNIRQLRAITIEHPORDFLSCLPV-----ELGKTLNLTGYDLKVAQV 150
QY 176 CKEWYRVTSQDMMLKKL-IERMV-----RTDSLMRGLAERRGMOYLFKNRPPDGNAP 227
Db 151 SKNMKLISEIKIMKSLGVEEFKHHPTDRTVGTAMGCTAIAAG-----VTIPDIIOP 203
QY 228 PN-SFYRALYPRKIIDITETISNMRGHRSLQRIHCHSETSKGYTL 273
Db 204 CDLVNHRFLKLOKFGDIFERRADKSRYLAKDKIERKNANPIMGSAV-LRGHEDHYTCM 262
QY 274 QYDDOKIVSGIRNDTIKIMDKNTLECKRIILTGHTSVYLCLOYDE--RVITGSSDSTVVR 331
Db 263 QIHDDVLTVYGSDDMTLAKWCIKDKGEVYTLVGHGTGWYTSQISCGKXIVSGSTDRIVKY 322
QY 332 WDVNTGEMNLTLIHCEAVLHLRFNNGMAYVTSKRSIAVWMDASPTDITLRLVGVHRA 391
Db 323 WSTVDSGLTLHTLOGHTSVRCMAMAGSILVYTGSHDITLRLVWDESGRHIA---TLGHHA 379
QY 392 AVNVVDFDKRIYASASDRTIKWNTSTCFVFTLNGHKGKGIACLOYRDR--LVYSSGSSD 449
Db 380 AVRQVQEDGDTIVVSGYDFTVKIMNAHTGRCIRTLGHNHNRVYSLFESRSIVSGSLD 439

Query Match 13.2%; Score 399; DB 4; Length 779;
 Best Local Similarity 24.3%; Pred. No. 3e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCLASTAKTEKNCVAKTKLANGTSSMIVPKOKRLASYTEK-----ELCVKFEQW 108
 D 194 TPLAKTTKTINN-----NNNIADLIESKDSIIISPEYLSDEIFSAIINNMLPAPYK-- 244
 QY 109 SESDOVEFEVHLISQWCHYOHINSYLKPMLODFITLAPARGIDHIAENILSYLDKAS 168
 D 245 -----NLFRIVANMDRELSDELGTILKDKRLITSLPF-----EISLKFNTLOFED 294
 QY 169 ICAAEIYCKEYR-VTSDGMLMKLI--ERMYRTDSLWGLAERBGWGYLFKNKPPDGN 225
 D 295 IINSLGVSQNMNKIIRKSTSLMKKLIISENFV-----SPKGF 331
 QY 226 APPNSEFYRALYPKIIOD-----IET--IESNMRCGRHSIORHCRSETSKGYCLOYD 276
 D 332 NSLNLKLSQKYPKLSOODRLRLSFLFNIFILKNWYKPVQRTTLRGHMTSVITCLOFE 391
 QY 277 DOKIVSGLRDNITIKINDKNTLECKRILHTGTSVLCLOYD--RVIIITGSSDSSTVAVMDVN 335
 D 392 DNYVITGADDKMIRYVDSINRKFLLQLSGHDGVAALKYAHGILVSGSDTRIVRWMDIK 451
 QY 336 TGEMLNTLIHCEAVLHLRFNNGMAYTCSKDRSIAYWDMASPTDITLRVLYGHRAAVNV 395
 D 452 KG-----CCT-----HVEEGHNSIVRC 468
 QY 396 VDDDD-----KIYVSASGDRTIKVNT-----STGE-----FVPTLNG 428
 D 469 LDIVYEKKIKIYVIGSRDNTLHWKLPKRESSVPDHGEHNDYPLVHTPEENPFVGYLNG 528
 QY 429 HKGIAIOLYDRILVYSSSDNTIRLMDIEGACLRVLEGHHELVRCIRPDN--KRIYSG 486
 D 529 HMASVFTVSGHGNIYVSSGYDNTLIVWDVAQMKCLYILSGHTRISTYIYHERRKRCISA 588
 QY 487 AYDGKIKVWL-----VAALDPRAPAGTL--CLRTLVEHSGRVRLQFDEQIVSS 535
 D 589 SMDTTIRIMDLLENIMNNGECYATNNSASPCAKIIGANTYTLQGHRTALVGLRLSDKFLVSA 648
 QY 536 SHDDTIILWDFLNDPAAQAEPPRSPSRITYY 566
 D 649 AADGSIRGMD-AND-----YSRKFYSY 668

RESULT 9

US-08-473-089-32

Sequence 32, Application US/08473089

Patent No. 6342368

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,089

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 779 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CD4 / CD20 protein, fig. 15
 US-08-473-089-32

Query Match 13.2%; Score 399; DB 4; Length 779;
 Best Local Similarity 24.3%; Pred. No. 3e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCLASTAKTEKNCVAKTKLANGTSSMIVPKOKRLASYTEK-----ELCVKFEQW 108
 D 194 TPLAKTTKTINN-----NNNIADLIESKDSIIISPEYLSDEIFSAIINNMLPAPYK-- 244
 QY 109 SESDOVEFEVHLISQWCHYOHINSYLKPMLODFITLAPARGIDHIAENILSYLDKAS 168
 D 245 -----NLFRIVANMDRELSDELGTILKDKRLITSLPF-----EISLKFNTLOFED 294
 QY 169 ICAAEIYCKEYR-VTSDGMLMKLI--ERMYRTDSLWGLAERBGWGYLFKNKPPDGN 225
 D 295 IINSLGVSQNMNKIIRKSTSLMKKLIISENFV-----SPKGF 331
 QY 226 APPNSEFYRALYPKIIOD-----IET--IESNMRCGRHSIORHCRSETSKGYCLOYD 276
 D 332 NSLNLKLSQKYPKLSOODRLRLSFLFNIFILKNWYKPVQRTTLRGHMTSVITCLOFE 391
 QY 277 DOKIVSGLRDNITIKINDKNTLECKRILHTGTSVLCLOYD--RVIIITGSSDSSTVAVMDVN 335
 D 392 DNYVITGADDKMIRYVDSINRKFLLQLSGHDGVAALKYAHGILVSGSDTRIVRWMDIK 451
 QY 336 TGEMLNTLIHCEAVLHLRFNNGMAYTCSKDRSIAYWDMASPTDITLRVLYGHRAAVNV 395
 D 452 KG-----CCT-----HVEEGHNSIVRC 468
 QY 396 VDDDD-----KIYVSASGDRTIKVNT-----STGE-----FVPTLNG 428
 D 469 LDIVYEKKIKIYVIGSRDNTLHWKLPKRESSVPDHGEHNDYPLVHTPEENPFVGYLNG 528
 QY 429 HKGIAIOLYDRILVYSSSDNTIRLMDIEGACLRVLEGHHELVRCIRPDN--KRIYSG 486
 D 529 HMASVFTVSGHGNIYVSSGYDNTLIVWDVAQMKCLYILSGHTRISTYIYHERRKRCISA 588
 QY 487 AYDGKIKVWL-----VAALDPRAPAGTL--CLRTLVEHSGRVRLQFDEQIVSS 535
 D 589 SMDTTIRIMDLLENIMNNGECYATNNSASPCAKIIGANTYTLQGHRTALVGLRLSDKFLVSA 648
 QY 536 SHDDTIILWDFLNDPAAQAEPPRSPSRITYY 566
 D 649 AADGSIRGMD-AND-----YSRKFYSY 668

RESULT 10

US-08-487-072A-32

Sequence 32, Application US/08487072A

Patent No. 6423684

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,072A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: MORASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 779 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15
 US-08-487-072A-32

Query Match 13.2%; Score 399; DB 4; Length 779;
 Best Local Similarity 24.3%; Pred. No. 3e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTEVCVAKTKLANGSSMIVPKORKLSASYEKER-----ELCVYFEQW 108
 194 TTPPLAKTTKTTNN-----NNNIADLESKDSITISPELSDIFSAINNHLPHAYFR-- 244
 109 SESDQVEFEVHLISOMCHYOHGHSYLYKPMLODFITALPARGLDHIAENILSYLDAKS 168
 245 -----NLFLRYANNDRELSIDGLIKONLKRDLITSLP-----EISLKFNYLOFED 294
 169 LCAAEVLCCKEYR-VTSDGLMKKLI--ERMVRTSLMGLAERGGQYLFKNKPPDGN 225
 295 IINSLGVSONMKNKTIKSTSLMKKLISENFV-----SPKGF 331
 226 APPNSFYRALYPKTIOD-----IET--IESNMRCGRHSLORIHCRSETSKGYCLOYD 276
 332 NSNLKLSOKYKPKLSQODRLRLSFLFNIFLKNMYPKFPVORTTLRGHMTSVITCLOPE 391
 277 DOKIVSGLRDNTIKIMDKNTLECKRILITGHTGSVLCLOYDE-RVLIITGSSDSTVRYWDVN 335
 392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGVMALKYAHGGLVSGSDTRVRYWDIK 451
 336 TGEMLNTLHHCBAVLHLRFNNGMAYTCSKDRSIAVMDMASPTDITLRYLVGHRAAVN 395
 452 KG-----CCT-----HYFEGHNSIVRC 468
 396 VDFDD---KYIVASGDRITIKVWNT-----STCE---FVRTLNG 428
 469 LDIVYEKNKIYIVTGSRDNTLHWKLPKSSVDPDGEHNDYPLVFRHTPEENRYFVGLRG 528
 429 HKRGIAQLQYDRLVYSGSSDNTIRLMDIEGACLVLEGEHELVCIRFDN--KRIYSG 486
 529 HMASVRYVSGHNIYVSGSDNTLIYWDVAQMKCLVILSGHTRIVSTIYDHERKRCISA 588
 487 AYDGIKIKVMDL-----VAALDPRAPAGTL--CLRTLVHSGVRFYLODFEFOIVSS 535

DB 589 SMDTIRIMDLLENINWNECSYATNSAPCAKILICAMTYTLOGHTALVGLRLSDKFLVSA 648
 536 SHDDTILIMDFLNDPAAQAEPPRSPRTTY 566
 DB 649 AADGSIRGMD-AND-----YSRKFY 668

RESULT 11
 US-09-177-165A-29
 Sequence 29, Application US/09177165A
 Patent No. 6426205
 GENERAL INFORMATION:
 APPLICANT: Tyers, Mike
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
 FILE REFERENCE: 11757, 10USU1
 CURRENT APPLICATION NUMBER: US/09/177,165A
 PRIOR FILING DATE: 1998-10-22
 PRIOR APPLICATION NUMBER: 60/092,443
 PRIOR FILING DATE: 1998-07-10
 PRIOR APPLICATION NUMBER: 60/063,254
 NUMBER OF SEQ ID NOS: 50
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 29
 LENGTH: 779
 TYPE: PRT
 ORGANISM: Saccharomyces cerevisiae
 US-09-177-165A-29

Query Match 13.2%; Score 399; DB 4; Length 779;
 Best Local Similarity 24.3%; Pred. No. 3e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTEVCVAKTKLANGSSMIVPKORKLSASYEKER-----ELCVYFEQW 108
 194 TTPPLAKTTKTTNN-----NNNIADLESKDSITISPELSDIFSAINNHLPHAYFR-- 244
 109 SESDQVEFEVHLISOMCHYOHGHSYLYKPMLODFITALPARGLDHIAENILSYLDAKS 168
 245 -----NLFLRYANNDRELSIDGLIKONLKRDLITSLP-----EISLKFNYLOFED 294
 169 LCAAEVLCCKEYR-VTSDGLMKKLI--ERMVRTSLMGLAERGGQYLFKNKPPDGN 225
 295 IINSLGVSONMKNKTIKSTSLMKKLISENFV-----SPKGF 331
 226 APPNSFYRALYPKTIOD-----IET--IESNMRCGRHSLORIHCRSETSKGYCLOYD 276
 332 NSNLKLSOKYKPKLSQODRLRLSFLFNIFLKNMYPKFPVORTTLRGHMTSVITCLOPE 391
 277 DOKIVSGLRDNTIKIMDKNTLECKRILITGHTGSVLCLOYDE-RVLIITGSSDSTVRYWDVN 335
 392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGVMALKYAHGGLVSGSDTRVRYWDIK 451
 336 TGEMLNTLHHCBAVLHLRFNNGMAYTCSKDRSIAVMDMASPTDITLRYLVGHRAAVN 395
 452 KG-----CCT-----HYFEGHNSIVRC 468
 396 VDFDD---KYIVASGDRITIKVWNT-----STCE---FVRTLNG 428
 469 LDIVYEKNKIYIVTGSRDNTLHWKLPKSSVDPDGEHNDYPLVFRHTPEENRYFVGLRG 528
 429 HKRGIAQLQYDRLVYSGSSDNTIRLMDIEGACLVLEGEHELVCIRFDN--KRIYSG 486
 529 HMASVRYVSGHNIYVSGSDNTLIYWDVAQMKCLVILSGHTRIVSTIYDHERKRCISA 588
 487 AYDGIKIKVMDL-----VAALDPRAPAGTL--CLRTLVHSGVRFYLODFEFOIVSS 535
 536 SHDDTILIMDFLNDPAAQAEPPRSPRTTY 566

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:54:27 ; Search time 39 Seconds

(without alignments)
3006.175 Million cell updates/sec

Title: US-09-601-168b-2

Perfect score: 3034

Sequence: 1 MDPAAVLQERALKFMNSE.....PAAQEPSPSPRTTYISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organellar:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriophage:*
17: sp-archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2997	98.8	569	11	09QUT5	09QUT5 mus musculus
2	2990	98.5	569	11	092159	092159 mus musculus
3	2978	98.2	569	11	092167	092167 mus musculus
4	2286	75.3	563	11	0923H0	0923H0 mus musculus
5	2089.5	68.9	510	5	044382	044382 drosophila
6	2089.5	68.9	510	5	09VDE3	09VDE3 drosophila
7	1498	49.4	430	5	09B554	09B554 heterodera
8	1230	40.5	252	11	0922C7	0922C7 mus musculus
9	661	21.8	1326	5	09VZF4	09VZF4 drosophila
10	640	21.1	353	4	09N0X6	09N0X6 homo sapien
11	640	21.1	561	4	096R12	096R12 homo sapien
12	640	21.1	589	4	096LE0	096LE0 homo sapien
13	640	21.1	627	4	096A16	096A16 homo sapien
14	640	21.1	707	4	0969H0	0969H0 homo sapien
15	639	21.1	629	4	08VHP4	08VHP4 mus musculus
16	639	21.1	629	11	08VBA4	08VBA4 mus musculus

17	638	21.0	122	13	08UUN3	08UUN3 xenopus lae
18	524	17.3	585	5	095ZT0	095ZT0 caenorhabdi
19	520	17.1	587	5	044083	044083 caenorhabdi
20	469	15.5	1376	3	08XIP2	08XIP2 podospora a
21	461.5	15.2	1356	3	08XIP4	08XIP4 podospora a
22	460.5	15.2	1356	3	08XIP5	08XIP5 podospora a
23	460.5	15.2	1356	3	08XIP3	08XIP3 podospora a
24	433.5	13.6	1227	16	0820R1	0820R1 anabaena sp
25	393	13.0	77	13	08UUN8	08UUN8 xenopus lae
26	387.5	12.8	304	11	09D4T2	09D4T2 mus musculus
27	378.5	12.5	1189	16	08YLO9	08YLO9 anabaena sp
28	375	12.4	1747	16	0820Z0	0820Z0 anabaena sp
29	365.5	12.0	942	5	096611	096611 dictyostell
30	361.5	11.9	677	16	08YDL9	08YDL9 anabaena sp
31	358.5	11.8	559	16	08YSC0	08YSC0 anabaena sp
32	357	11.8	410	13	090ZL4	090ZL4 xenopus lae
33	356	11.7	410	13	09PFR5	09PFR5 gallus gall
34	354	11.7	410	4	08WZ89	08WZ89 homo sapien
35	351	11.6	411	6	09GL51	09GL51 sus scrofa
36	344	11.3	411	5	096698	096698 drosophila
37	340	11.2	410	4	08WZ88	08WZ88 homo sapien
38	339.5	11.2	777	3	09USN3	09USN3 schizosacch
39	336	11.1	676	16	08YSG6	08YSG6 anabaena sp
40	335.5	11.1	353	5	08T776	08T776 branchiosto
41	334.5	11.0	283	11	093CV3	093CV3 mus musculus
42	334.5	11.0	454	4	08WV51	08WV51 homo sapien
43	334.5	11.0	454	4	09HA09	09HA09 homo sapien
44	334.5	11.0	594	11	0922B6	0922B6 mus musculus
45	334	11.0	589	16	08YV14	08YV14 anabaena sp

ALIGNMENTS

RESULT 1
ID 09QUT5 PRELIMINARY; PRT; 569 AA.
AC 09QUT5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)
DE (F-box WD40 repeat protein 1).
GN BTRC OR FBXW1.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99199275; PubMed=10097128;
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hattori K., Higashi H., Nakano H., Okumura K., Ono K., Good R.A.,
RA Nakayama K.-i.;
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT ubiquitin ligase Skp1/Cul1/F-box protein FWD1."
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase."
RL Nature 396:590-594(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX Strausberg R.;
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=129/SV;
RX MEDLINE=21601157; PubMed=11735228;

RA Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,
 RA Nakayama K., I.
 RT "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of
 RT Caenorhabditis elegans SEL-10."
 RL Genomics 78:214-222(2001).
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: AF081887; AAD17755.1; -;
 DR EMBL: AF099332; AAD08701.1; -;
 DR EMBL: BC003989; AAO03989.1; -;
 DR EMBL: AF391190; AAL40929.1; -;
 DR EMBL: AF391178; AAL40929.1; JOINED.
 DR EMBL: AF391179; AAL40929.1; JOINED.
 DR EMBL: AF391180; AAL40929.1; JOINED.
 DR EMBL: AF391181; AAL40929.1; JOINED.
 DR EMBL: AF391182; AAL40929.1; JOINED.
 DR EMBL: AF391183; AAL40929.1; JOINED.
 DR EMBL: AF391184; AAL40929.1; JOINED.
 DR EMBL: AF391185; AAL40929.1; JOINED.
 DR EMBL: AF391186; AAL40929.1; JOINED.
 DR EMBL: AF391187; AAL40929.1; JOINED.
 DR EMBL: AF391188; AAL40929.1; JOINED.
 DR EMBL: AF391189; AAL40929.1; JOINED.
 DR MGD: MGI:1338871; Btrc.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR ProDom: PD000018; WD40; 4.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ligase: Repeat; WD repeat.
 SQ SEQUENCE 569 AA; 65105 MM; BCTD6544815B2296 CRC64;

Query Match 98.8%; Score 2997; DB 11; Length 569;
 Best Local Similarity 98.6%; Pred. No. 7.1e-253;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAEAVLQERKALFKPMSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLA 60
 DB 1 MDPAEAVLQERKALFKPMSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLA 60
 QY 61 STAKTEKNCVAKTKLANGTSSMIVPKORKLSASYEKKEKLCVYFEOMSESDDVEFEHL 120
 DB 61 STAKTEKNCVAKTKLANGTSSMIVPKORKLSASYEKKEKLCVYFEOMSESDDVEFEHL 120
 QY 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 DB 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 QY 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 DB 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 QY 181 RYVSDGMLMKKLIERNVRTDSIMRGLAERGGWGYLFKNRPDENAPNSFYALYPKII 240
 DB 181 RYVSDGMLMKKLIERNVRTDSIMRGLAERGGWGYLFKNRPDENAPNSFYALYPKII 240
 QY 241 QDIETIESNMRGSHSLQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 QDIETIESNMRGSHSLQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGVSILCLQYDERVITIGSSDSIVRVWDVNTGEMLNTLIHHCCEAVLHLRFNNGM 360
 DB 301 RILTGHTGVSILCLQYDERVITIGSSDSIVRVWDVNTGEMLNTLIHHCCEAVLHLRFNNGM 360
 QY 361 VYCSKDRSLAVWDMASPTDITLRLVYVGHRAAVNVVDEPDXYIVASGDRITKIVNTSTIC 420
 DB 361 VYCSKDRSLAVWDMASPTDITLRLVYVGHRAAVNVVDEPDXYIVASGDRITKIVNTSTIC 420
 QY 421 EEFVRLNGHKGRIACLOYBDRLVVSGSSDNTIRLWDIEGACLRVLEGHEELVRCIRFDN 480
 DB 421 EEFVRLNGHKGRIACLOYBDRLVVSGSSDNTIRLWDIEGACLRVLEGHEELVRCIRFDN 480

QY 481 KRIVSAYDYGKIKVMDVLAADPPRAPAGTLCRLTVHSGRVRLQDFEQIVSSSHDT 540
 DB 481 KRIVSAYDYGKIKVMDVLAADPPRAPAGTLCRLTVHSGRVRLQDFEQIVSSSHDT 540
 QY 541 IILWDFLNDPAAPAEPPSPRSRTYIYSR 569
 DB 541 IILWDFLNDPAAPAEPPSPRSRTYIYSR 569

RESULT 2
 ID 092159 PRELIMINARY; PRT; 569 AA.
 AC 092159;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-2002 (TREMBLrel. 20, Last annotation update)
 DE Beta-transducin repeat containing protein.
 GN BTRC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N. A.
 RX MEDLINE=99145465; PubMed=9990853;
 RA Spencer E., Jiang J., Chen Z.J.;
 RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
 RT Slimb/Delta-TICP."
 RL Genes Dev. 13:284-294(1999).
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: AF112979; AAD04181.1; -;
 DR MGD: MGI:1338871; Btrc.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR ProDom: PD000018; WD40; 4.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 569 AA; 65047 MM; BCTC7A44815BED96 CRC64;

Query Match 98.5%; Score 2990; DB 11; Length 569;
 Best Local Similarity 98.4%; Pred. No. 2.9e-252;
 Matches 560; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDPAEAVLQERKALFKPMSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLA 60
 DB 1 MDPAEAVLQERKALFKPMSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLA 60
 QY 61 STAKTEKNCVAKTKLANGTSSMIVPKORKLSASYEKKEKLCVYFEOMSESDDVEFEHL 120
 DB 61 STAKTEKNCVAKTKLANGTSSMIVPKORKLSASYEKKEKLCVYFEOMSESDDVEFEHL 120
 QY 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 DB 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 QY 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 DB 121 ISQCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 QY 181 RYVSDGMLMKKLIERNVRTDSIMRGLAERGGWGYLFKNRPDENAPNSFYALYPKII 240
 DB 181 RYVSDGMLMKKLIERNVRTDSIMRGLAERGGWGYLFKNRPDENAPNSFYALYPKII 240
 QY 241 QDIETIESNMRGSHSLQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 QDIETIESNMRGSHSLQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGVSILCLQYDERVITIGSSDSIVRVWDVNTGEMLNTLIHHCCEAVLHLRFNNGM 360
 DB 301 RILTGHTGVSILCLQYDERVITIGSSDSIVRVWDVNTGEMLNTLIHHCCEAVLHLRFNNGM 360

```

Db 301 RILGHTGSLVLCQYDERVITIGSSDSTVRWVDVNAEMINTLIHCEAVLHLRFNNMGM 360
Qy 361 VTCGSRSLAVWMDASPDITLRLVGVHRAAVNVDPDDKYIVASGDRITIKVWNTSTC 420
Db 361 VTCGSRSLAVWMDASPDITLRLVGVHRAAVNVDPDDKYIVASGDRITIKVWNTSTC 420
Qy 421 EFVRLNGHRRGIACQYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
Db 421 EFVRLNGHRRGIACQYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
Qy 481 KRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTVHSGRVFLQDFEFQIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTVHSGRVFLQDFEFQIVSSSHDT 540
Qy 541 ILIMDFLNDPAQAAPPSPRTTYTISR 569
Db 541 ILIMDFLNDPAQAAPPSPRTTYTISR 569

```

RESULT 3

09RIG7 PRELIMINARY; PRT; 569 AA.

```

AC 09RIG7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Beta-transducin repeat-containing protein.
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Winston J., Elledge S.J., Harper J.W.;
RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF110396; AAD41025.1; -.
DR MGD: MGI:133871; Btrc.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPOTELINRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

```

Query Match 98.2%; Score 2978; DB 11; Length 569;
 Best Local Similarity 97.9%; Pred. No. 3.2e-251;
 Matches 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

Qy 1 MDPAEAVLQEKALKFNSSREDCNNGEPKRIIPKNSLRQYNSCARLINOETVCLA 60
Db 1 MDPAEAVLQEKALKFNSSREDCNNGEPKRIIPKNSLRQYNSCARLINOETVCLA 60
Qy 61 STAKTECVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFEQWSESDQVEFEHL 120
Db 61 STAKTECVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFEQWSESDQVEFEHL 120
Qy 121 ISQCHYOHGHINSLYKMLQDPTITLPAAGLDHIAENIISYDAKSLCAELVCWEKY 180
Db 121 ISQCHYOHGHINSLYKMLQDPTITLPAAGLDHIAENIISYDAKSLCAELVCWEKY 180
Qy 181 RYSDGMVKKLIERMVRTDSLMRGLAERRGQGYLTKNNPPDGNAPNSRYALYPKII 240
Db 181 RYSDGMVKKLIERMVRTDSLMRGLAERRGQGYLTKNNPPDGNAPNSRYALYPKII 240

```

```

Db 181 RYSDGMVKKLIERMVRTDSLMRGLAERRGQGYLTKNNPPDGNAPNSRYALYPKII 240
Qy 241 ODIEETESNMRCGRHSIORICRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDNTECK 300
Db 241 ODIEETESNMRCGRHSIORICRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDNTECK 300
Qy 301 RILGHTGSLVLCQYDERVITIGSSDSTVRWVDVNAEMINTLIHCEAVLHLRFNNMGM 360
Db 301 RILGHTGSLVLCQYDERVITIGSSDSTVRWVDVNAEMINTLIHCEAVLHLRFNNMGM 360
Qy 361 VTCGSRSLAVWMDASPDITLRLVGVHRAAVNVDPDDKYIVASGDRITIKVWNTSTC 420
Db 361 VTCGSRSLAVWMDASPDITLRLVGVHRAAVNVDPDDKYIVASGDRITIKVWNTSTC 420
Qy 421 EFVRLNGHRRGIACQYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
Db 421 EFVRLNGHRRGIACQYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
Qy 481 KRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTVHSGRVFLQDFEFQIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTVHSGRVFLQDFEFQIVSSSHDT 540
Qy 541 ILIMDFLNDPAQAAPPSPRTTYTISR 569
Db 541 ILIMDFLNDPAQAAPPSPRTTYTISR 569

```

RESULT 4

0923HO PRELIMINARY; PRT; 563 AA.

```

AC 0923HO;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE F-box/WD40 repeat-containing protein HOS.
GN FBXW1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=FVB/N;
RA Bhatia N., Hertler J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;
RT "Mouse homolog of HOS (mHOS) is overexpressed in skin tumors and
RT implicated in constitutive activation of NF-kappaB."
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AT038079; AAK72095.1; -.
DR MGD: MGI:2144023; Fbxw1b.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRODOM: PD000018; WD40; 4.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 563 AA; 64741 MW; 9AB562F3FF5E3496 CRC64;

```

Query Match 75.3%; Score 2286; DB 11; Length 563;
 Best Local Similarity 75.7%; Pred. No. 8e-191;
 Matches 440; Conservative 46; Mismatches 65; Indels 30; Gaps 7;

```

Qy 1 MDPAEAVLQEKALKFNSSREDCNNGEPKRIIPKNSLRQYNSCARLINOETVCLA 60
Db 1 MDPAEAVLQEKALKFNSSREDCNNGEPKRIIPKNSLRQYNSCARLINOETVCLA 60
Qy 56 TVCLASTA-MKTEN-----CVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFEQW 108
Db 56 TVCLASTA-MKTEN-----CVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFEQW 108
Qy 45 VRLQNTSVWEDQNEDESPKKSALMQLISNCTSSVITSRRKPSBGNQKQKDCICITYFDQM 104
Db 45 VRLQNTSVWEDQNEDESPKKSALMQLISNCTSSVITSRRKPSBGNQKQKDCICITYFDQM 104

```

	Best Local Similarity	78.3% ; Pred.	No. 9_8e-174;	Matches	394; Conservative	48; Indels	54; Gaps	7;
OY	64 MTEVCVAKTKLANG---	TSSMIVPQKRLSAS--YEKEELCVKYFEQMSESDOYEVE	118					
Dd	4 METDKIMDET-N-SMAQAFVTMTMLDYVRKKDSSPYQTRELCFOFYQWMSBGQDFVE	62						
OY	119 HHSQCCHQHGHINSYLPMLORDITPLPAQRGDHIAENLSTIYLDASLCAAEUYCKE	178						
Dd	63 HLISRCHQHGHQINAYLKPMLORDITPLPIKGIDHIENLTSLYDAESLSSELLEYCE	122						
OY	179 WTVATSDGMAMKLIERAVRTDSLNRGLAERGWGQYLEKKNPPDGNAFNSFYRALYPK	238						
Dd	123 WLRIVISEGLMKLELRKVRTDSLNRGLAERRNNMQYLEFKPRPG-QTORPHSFHRLEPK	181						
OY	239 IIODIETTESNRCGHSHLORHCSESYSKYCYCLOYDOQAITYSGIRONTIYMKNILE	298						
Dd	182 IMNDIDSIEENNRRGTGHMLRINCENSENKGYCCLOYDGKIVSGLRDNWKIKIMVRTDLO	241						
OY	299 CKRIITLGHTGSVLICLOYDERVIITITSSDSYVVMVOVNTGEMINTLIHHCEAVLHPFNNG	358						
Dd	242 CVKTLMGHTGSVLICLOYDYDKVIASSDSYSTVAVDYNTEGANVTLLIHCEAVLHLFPNG	301						
OY	339 MWVTSCSKDRSIADVMDASFDTILTRRVLYGNRAAVNVDDFDKYTVSAGDRTIKWMNTS	418						
Dd	302 MWVTSCKRSRIAVMPMTSPSETTLRRVLYGNRAAAVVNVDFDEKYTVSAGDRTIKWMSFS	361						
OY	419 TCFEFTNLNGHGRIACLOYRDRLVVGSSDNTITLMIDEGCACLRVLEGHELVRCIRF	478						
Dd	362 TCFEFTNLNGHGRIACLOYRDRLVVGSSDNSIKLMIIEGCACLRVLEGHELVRCIRF	421						
OY	479 DNKRIVSGADGIKIWMDLVAALDPAPADGCTCLTIVEHSGRVRLTFDEFQIYSSHD	538						
Dd	422 DTKRIVSAGADGIKIWMDLVAALDPRASNLTCLNTVTEHTGRVERLGFDEFQIYSSHD	481						
OY	539 DTILLWDLPNDPAQAEPFRPS	561						
Dd	482 DTILLWDLPNTPNENKTGRTPS	504						
	RESULT 6							
ID	09YDE3	PRELIMITARY:	PRT:	510 AA.				
AC	09YDE3.							
DT	01-MAY-2000 (TREMBLrel. 13, Created)							
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)							
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)							
DE	SIMB protein (SIMB).							
GN	SLMB OR CG3412.							
OS	Drosophila melanogaster (fruit fly).							
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;							
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;							
OC	Ephydroidea; Drosophilidae; Drosophila.							
OX	NCHI_Taxid=7227;							
NN	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN-BERKELEY;							
RX	MEDLINE=20196006; PubMed=10731132;							
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,							
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,							
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,							
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,							
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,							
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,							
RA	Abdil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,							
RA	Ballew R.M., Basay A., Baxendale J., Bayraktaroglu L., Beasley E.M.,							
RA	Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,							
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,							
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,							
RA	Cherry J.M., Casley S., Dahlke C., Davenport L.B., Davies P.,							
RA	de Paolis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,							
RA	Dodson K.J., Doull L.E., Downes M., Duran-Rocha S., Dunkov B.C., Dunn P.,							
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,							
RA	Lurie D.I., Wang X., Yin J., Zhuo R., Smith R.O., Wilson R.,							
RA	Zhang J., White O., Adams M.D., Kerlavage A.H., Sutton G., Olsen I.,							
RA	McClure M., Venter A., Orkin R.C., Green E.D., McPherson J.R.,							
RA	Karp P., Kim U., Pruitt K., Fritch J., Davis R., Shinnar R.,							
RA	Hillier B.G., Birney D., Scott A., Weissenbach J., Young R.A.,							
RA	Carmelli D., Frith C.E., Miller W., Jones R., Berry A., Mehlhorn A.,							
RA	Gravert C., Gunter T., Reuter T., Fulton R.S., Staden R., Steward R.,							
RA	Stumpel C., Schein J., Wei X., Wu J., Yan H., Zhao Y., Zheng X.,							
RA	Miller W., Berg D., Armstrong J., Haussler D., Langdon C.B., Lin M.,							
RA	Roche B.,							


```

Db 241 ASTCEVFTLNGKRGKGIACLOVRDLIVSGSSDYSIRLMDIEGNCGLVLEHGEIYWC 300
QY 477 RPDNRRIYSGAVDGKIKWDLVAALDPAPAGTLCRLTVHSGRVFLQDFEQIYSS 536
Db 301 RFDNRRIYSGAVDGKIKWDLVAALDPAPAGTLCRLTVHSGRVFLQDFEQIYSS 360
QY 537 HDTLLIMDFLNP-DEPA-----CAEPPRSPS 561
Db 361 HDTLLIMDFLNPGPSAAVAAAGAAAAAHNAQVQHQAAPAAQPA 405

RESULT 8
Q922C7 PRELIMINARY; PRT; 252 AA.
ID 0922C7
AC 0922C7;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Similar to f-box and WD-40 domain protein 1B (Fragment).
GN FBXW1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; BC008552; AA08552.1; -.
DR MGI; MGI:2144023; Fbxw1b.
DR InterPro; IPR001680; WD40.
DR Pfam; PF004400; WD40; 6.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 3.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE; PS50082; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT NON_TER
SQ SEQUENCE 252 AA; 28424 MW; F71737CBD7A9F75F CRC64;

Query Match 40.5% Score 1230; DB 11; Length 252;
Best Local Similarity 92.9% Pred. No. 3.6e-99;
Matches 234; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 318 RVTITGSSDSTVAVWVNTGEMTLTIHCEAVLHFRNNGMAYTCKSDSIWVMDASP 377
Db 1 RVTITGSSDSTVAVWVNTGEMTLTIHCEAVLHFRNNGMAYTCKSDSIWVMDASA 60
QY 378 TDTLRRLVGHRAAVNVDFDQKYIVSASGDRTIKMYNTSTCEVFTLNGKRGKGIACLO 437
Db 61 TDTLRRLVGHRAAVNVDFDQKYIVSASGDRTIKMYNTSTCEVFTLNGKRGKGIACLO 120
QY 438 YRDLVAVSGSSDNTIRLMDIEGACLVLEHGEIYWCIRPDNRRIYSGAVDGKIKWDL 497
Db 121 YRDLVAVSGSSDNTIRLMDIEGACLVLEHGEIYWCIRPDNRRIYSGAVDGKIKWDL 180
QY 498 VVALDPAPAGTLCRLTVHSGRVFLQDFEQIYSSHDTLLIMDFLNPDAAPAEPP 557
Db 181 QALDPAPAGTLCRLTVHSGRVFLQDFEQIYSSHDTLLIMDFLNPDAAPAEPP 240
QY 558 RSPSRITTYISR 569
Db 241 RSPSRITTYISR 252

RESULT 9
Q9VZF4 PRELIMINARY; PRT; 1326 AA.
ID 09VZF4;
AC 09VZF4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

```

```

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE CG15010 protein (LD21322p) (LD30271p).
GN A60 OR CG15010.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Ffankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos A.C., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spidling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Frise E., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AE003480; AAG22246.1; -.
DR EMBL; AY061300; AAL28848.1; -.
DR EMBL; AY075401; AAL68231.1; -.

```


OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCHI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=2003061; PubMed=10531037;
 RA Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.,
 RT "A family of mammalian F-box proteins.";
 RL Curr. Biol. 9:1180-1182(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Koepf D.M., Winston J.T., Harper J.W., Ellledge S.J.,
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: AF383178; AAK60269.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR ProDom: PD000018; WD40; 2.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 DR Repeat; WD repeat.
 KM NON-TER 1
 FT SEQUENCE 561 AA; 63165 MW; B81C1E2206B0D88 CRC64;

Query Match 21.1%; Score 640; DB 4; Length 561;
 Best local similarity 30.8%; Pred. No. 3,7e-47;
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KPNNSEREDCNNGEPKRIPEKNSLQTYNSCARLCINQETVCLASTAMKTENCYAKT 73
 Db KIDHGEVRSFSLGKPKCKV-----SEYTSYTGTL-----VPCSA-----TPPTFGDL 66
 QY 74 KLANGSSMIYPRKORLSYEKEKEL--CYKFEOMSESDOVEFEHLISOMCHYOGH 131
 Db 67 RAANGG-----QQRRTTSVQPTGLOEWMKMFQMSGPEKLALDELIDSCPTQVKA 121
 QY 132 INSYLKPMLQRFITAPARGIDHIAENILSYLDAKSLCAELVCKEYRATSDGMIMK 191
 Db 122 MQVIEPQFQRFISILP-----KELALYVLSFLEPKDLQAQTCRYRRIADENILMRE 177
 QY 192 LIERVWRITDSLMRGLAE-----RRCWGQYLKKNKPPGNAAPPNSFYALYPKIIQDIETI 246
 Db 178 KCKE-----EGIDEPHLIKRRK-----VIKPGFIHSPKMSAY-----IRQ--HRI 215
 QY 247 ESNMRCGRHSIQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRILITGH 306
 Db 216 DTMNRGELKSPKV-LKQHDHVTICLQPCGNRIYVSGSDNTLKWSAVTGKCLRTLVGH 274
 QY 307 TGSVLCLOYDERVITITGSSDSTVRYVWVNGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366
 Db 275 TGGVSSQMRNIIISGSTDRTLKVMNAETGECIHTLYGHTSVRCMHLHEKRYVSGSRD 334
 QY 367 RSIAYWMAASPDTILTRVLVGHRAAVNVDPDKIYVSASGDTIYVWNTSTICEVYRTL 426
 Db 335 ATLRYWDLDTGQCL---HYLMGHAAVAVCYQYDGRRVASGAYDPMVWMPETETCLHTL 391
 QY 427 NGHKRGIALCYRDLRYVSSGSDNTIRLMDIEGACLRVLEGEHELVRICIRFDNKRIVSG 486
 Db 392 QGHTNRVYSLQFDGIHVYSGSLDTSIRVMDVETGNCIHTLITGQSLTSGMELKDNILVSG 451
 QY 487 AYDGKIKYWDLVAAIDPRAPAGTICLRTLV---EHSGRVFRLOFDEFQIVSSSHDPTILT 543
 Db 452 NADSTVAKIMDIKTG-----QCLOTLOGPNKHOSAVTCLQFNKNFVITSSDDGTIVKL 502
 QY 544 WD 545
 Db 503 WD 504

RESULT 12
 ID 096LEO PRELIMINARY; PRT; 589 AA.
 AC 096LEO;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE F-box protein SEL10.
 GN SEL10.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCHI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Li J., Pauley A.M., Myers R.L., Shuang R., Bashler J.R., Yan R.,
 RA Buhl A.E., Gurney M.E.,
 RT "SEL-10 interacts with Presenilin 1, facilitates its ubiquitination,
 RT and Alters A-beta production.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: AV008274; AAG16640.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR ProDom: PD000018; WD40; 2.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 DR Repeat; WD repeat.
 KM NON-TER 1
 FT SEQUENCE 589 AA; 66120 MW; 2AFB6E8A36E8DE CRC64;

Query Match 21.1%; Score 640; DB 4; Length 589;
 Best local similarity 30.8%; Pred. No. 4e-47;
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KPNNSEREDCNNGEPKRIPEKNSLQTYNSCARLCINQETVCLASTAMKTENCYAKT 73
 Db KIDHGEVRSFSLGKPKCKV-----SEYTSYTGTL-----VPCSA-----TPPTFGDL 66
 QY 74 KLANGSSMIYPRKORLSYEKEKEL--CYKFEOMSESDOVEFEHLISOMCHYOGH 131
 Db 67 RAANGG-----QQRRTTSVQPTGLOEWMKMFQMSGPEKLALDELIDSCPTQVKA 121
 QY 132 INSYLKPMLQRFITAPARGIDHIAENILSYLDAKSLCAELVCKEYRATSDGMIMK 191
 Db 122 MQVIEPQFQRFISILP-----KELALYVLSFLEPKDLQAQTCRYRRIADENILMRE 177
 QY 192 LIERVWRITDSLMRGLAE-----RRCWGQYLKKNKPPGNAAPPNSFYALYPKIIQDIETI 246
 Db 178 KCKE-----EGIDEPHLIKRRK-----VIKPGFIHSPKMSAY-----IRQ--HRI 215
 QY 247 ESNMRCGRHSIQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRILITGH 306
 Db 216 DTMNRGELKSPKV-LKQHDHVTICLQPCGNRIYVSGSDNTLKWSAVTGKCLRTLVGH 274
 QY 307 TGSVLCLOYDERVITITGSSDSTVRYVWVNGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366
 Db 275 TGGVSSQMRNIIISGSTDRTLKVMNAETGECIHTLYGHTSVRCMHLHEKRYVSGSRD 334
 QY 367 RSIAYWMAASPDTILTRVLVGHRAAVNVDPDKIYVSASGDTIYVWNTSTICEVYRTL 426
 Db 335 ATLRYWDLDTGQCL---HYLMGHAAVAVCYQYDGRRVASGAYDPMVWMPETETCLHTL 391
 QY 427 NGHKRGIALCYRDLRYVSSGSDNTIRLMDIEGACLRVLEGEHELVRICIRFDNKRIVSG 486
 Db 392 QGHTNRVYSLQFDGIHVYSGSLDTSIRVMDVETGNCIHTLITGQSLTSGMELKDNILVSG 451
 QY 487 AYDGKIKYWDLVAAIDPRAPAGTICLRTLV---EHSGRVFRLOFDEFQIVSSSHDPTILT 543
 Db 452 NADSTVAKIMDIKTG-----QCLOTLOGPNKHOSAVTCLQFNKNFVITSSDDGTIVKL 502
 QY 544 WD 545
 Db 503 WD 504

```

OY 544 WD 545
DB 531 WD 532

RESULT 13
O96A16 PRELIMINARY; PRT; 627 AA.
ID O96A16:
AC O96A16:
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DE Archipelago beta form (F-box protein FBW7).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21449047; PubMed=11565033;
RA Moberg K.H., Bell D.W., Maher D.C., Haber D.A., Hariharan I.K.;
RT "Archipelago regulates cyclin E levels in Drosophila and is mutated in
RT human cancer cell lines.";
RL Nature 413:311-316(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=20003061; PubMed=10531037;
RA Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.;
RT "A family of mammalian F-box proteins.";
RL Curr. Biol. 9:1180-1182(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Koepf D.M., Winston J.T., Harper J.W., Ellledge S.J.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF411972; AAL06291.1; -
DR EMBL: AY033553; AAK57547.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 2.
DR ProDom: PD000018; WD40; 2.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 627 AA; 70324 MW; 3D4107C05381BED CRC64;

Query Match 21.1%; Score 640; DB 4; Length 627;
Best Local Similarity 30.8%; Pred. No. 4, 4e-47;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

OY 14 KFMNSSEEDCNGEPPRKIIPEKNSLRQYNSCARLCLNOETVCLASTAMKTEVCYAKT 73
DB 91 KLDGSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPTFGLD 132
OY 74 KLAGTSMYIPKORXLSASEKEKEL--CVKYPEOMSESDOYEFVHLISQCHYQHG 131
DB 133 RAAAGGOC-----QORRRITTSVOPPTGLOEWKMFQSWSGPEKLALDELIDSCPTPVKH 187
OY 132 INSYLKPMLORFETLAPARGLDIAENILSYLDAKSLCAELCKEYEVTSQGLMK 191
DB 188 MMGYIEPFOFORDFISLPL-----KELALYVLSFLEPKDLQAQYCRWRRLAEENLIMRE 243
OY 192 LIEHNVRTDSLMKGLAE-----RKGQGYLFKNKPPDGNAPPNPSFYALPKIIQDIETI 246
DB 244 KCKE-----EGIDPLHLIKRRK-----VIKPGFIHSPKSAV-----IRQ--HRI 281
OY 247 ESNRRCGHSRLDRHCHSETSKGVYCLQYDDOKITVSGLRDNTIKWKNTLECKRIITLGH 306
DB 282 DTNMRGELSPKV-LKGHDDVITTCLOFCGNRIVSGSDNTILKWSAVGICLKLITLVGH 340

```

```

OY 307 TGSVLCLQYDERVIITGSSDSTVRAWVDVNTGEMLTLLIHCEAVLHLRFNNGMAYTCSKD 366
DB 341 TGVWSSQMRDNIIISGSTDFLKLWNAETGECIHTLYGHSTVACMHLHKKRVSGSRD 400
OY 367 RSLAVWDMASPDITILRLRVLGHRAAVNVVDPDDKIYISASDRKIKVWNSTCFVPTL 426
DB 401 ATLRLWMDLETGGCL---HLMLGHVAAVRCVOYDGRVRSYGAIDFMVKWMDPETETCLHTL 457
OY 427 NGHKRGIACTQYRDLRVVSGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDNRKRVISG 486
DB 458 OCHTRNVYSLOPDGTHVYVSGSLDTSIRWVDVETGNCIHTLGHOSLTSGMELKDNIIIVSG 517
OY 487 AYDGKIKWDLVAALDPRAPACTLCLRLV---EHSGRVFLQDFEQIVSSSHDDTILI 543
DB 518 NDSVTYKIMDIKTG-----QCLQTLQPGPNKHQSANTCLOFKNKRVITSSDDGTVKL 568
OY 544 WD 545
DB 569 WD 570

```

```

RESULT 14
O969H0 PRELIMINARY; PRT; 707 AA.
ID O969H0:
AC O969H0:
DT 01-DEC-2001 (TREMblrel. 19, created)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE F-box protein CDC4 (Archipelago alpha form).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21449048; PubMed=11565034;
RA Strichmaler H., Spruck C.H., Kaiser P., Won K.A., Sangfelt O.,
RA Reed S.I.;
RT "Human F-box protein hcdc4 targets cyclin E for proteolysis and is
RT mutated in a breast cancer cell line.";
RL Nature 413:316-322(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=21449047; PubMed=11565033;
RA Moberg K.H., Bell D.W., Maher D.C., Haber D.A., Hariharan I.K.;
RT "Archipelago regulates cyclin E levels in Drosophila and is mutated in
RT human cancer cell lines.";
RL Nature 413:311-316(2001).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AY049984; AAL07271.1; -
DR EMBL: AF411971; AAL06290.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 2.
DR ProDom: PD000018; WD40; 2.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 707 AA; 79663 MW; EA4357F76D8203 CRC64;

Query Match 21.1%; Score 640; DB 4; Length 707;
Best Local Similarity 30.8%; Pred. No. 5, 2e-47;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

OY 14 KFMNSSEEDCNGEPPRKIIPEKNSLRQYNSCARLCLNOETVCLASTAMKTEVCYAKT 73
DB 171 KLDGSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPTFGLD 212

```

```

QY 74 KLANGSSMIVPORKORLSASYEKEL--CYKPFQMSSESDQVEVEHLISOMCHYOGH 131
Db 213 RAANGGQ-----QQRRTTSVOPPTGLOEMLKMFQMSGPEKLALDELIDSCPTQVKH 267
QY 132 INSYLKPMLQDRFETALPARGLDIAENILSYLDKSLCAAEVLCYKEMRYVTSQGMKMK 191
Db 268 MMQVTEPQFQDFISLIP-----KELALYVLSFLEPKDLQAQOTCRWRIIAEDNLMRE 323
QY 192 LIEMWRTSLMRGLAE-----RRCMGQYLFKNRPPDGNAPPNSTYRALLYPKIIODIETI 246
Db 324 KCKE-----EGIDEPHLIKRRK-----VTKPGFIHSPKMSAY-----IRO--HRI 361
QY 247 ESNMRCGRHSIORICRSETSKGYCLOYDDOKIYSGLRDNTIKIMDKNTLECKRIILGH 306
Db 362 DNNMRGELKSPKV-LKGHDHVITCLOFGNRIYSGSDNTLKYMSAVTGKCLRTLVGH 420
QY 307 TGSVLCLOYDERVIITSSSDSTVRYWVDNTGEMTLIIHCEAVLHLRFNNGMMVYCSKD 366
Db 421 TGGVWSSQMRDNIISGSTDTLTKVWNAETGECIHTLYGHTSVRCMHLHEKRYVSGSRD 480
QY 367 RSLAVWMDASPTDITLRVYLGHRRAAVNVDPDDKXIYASGDRITIKVWNTSTCEPVRTL 426
Db 481 ATLWMDIETGQCL--HYLMGHVAAVRCVQYDGRVVSAGVDFMKVWDEPTEETCLHTL 537
QY 427 NGHKRGIACTQYRDLRYVSSSDNTIRLMDIEGACILRVLEGEHELIVRCIRFNDKRIYSG 486
Db 538 QGHTNRVYSLQFDGIHVYSGSLDTSIRWVDVETGNCIHTLGHQSLTSGMELKONILVSG 597
QY 487 AYDGKIKVMDVLAALDPRAAGTLCRLTY---EHSGRVFRLOPDEFQIVSSSHDDTILI 543
Db 598 NADSTVKIMDIKTG-----QCLQTLQGPKNHOSAVTCLQFNKNFVITSSDDGTIVKL 648
QY 544 WD 545
Db 649 WD 650

```

RESULT 15

Q8VHP4 PRELIMINARY; PRT; 629 AA.

AC Q8VHP4;

DT 01-MAR-2002 (TReMBLrel. 20, Created)

DT 01-MAR-2002 (TReMBLrel. 20, last sequence update)

DE F-box protein.

GN FBWD6.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NBL_TaxID-10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Ilyin G.P.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF427101; AAL50052.1; -

DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.

DR Pfam: PF00646; F-box; 1.

DR PRINTS: PRO0320; GPROTEINRPT.

DR ProDom: PD000018; WD40; 2.

DR SMART: SM00256; PROX; 1.

DR SMART: SM00320; WD40; 7.

DR PROSITE: PS50181; FBOX; 1.

DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.

DR PROSITE: PS50082; WD_REPEATS_2; 7.

DR PROSITE: PS50294; WD_REPEATS_REGION; 1.

SO SEQUENCE 629 AA; 70562 MW; BE916405A3490A3E CRC64;

Query Match 21.1%; Score 639; DB 11; Length 629;

Best Local Similarity 30.8%; Pred. No. 5,4e-47;

Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```

QY 14 KEMNSEREDCNGNEPPRKIIPEKNSLRQTYNSCARLCLNDETCLASTAMKTENCVAKT 73
Db 93 KLDHGSEVASFSSIGKKPKV-----SDYTSITGL-----VPCSA-----TPTFQDL 134
QY 74 KLANGSSMIVPORKORLSASYEKEL--CYKPFQMSSESDQVEVEHLISOMCHYOGH 131
Db 135 RAANGGQ-----QQRRTTSVOPPTGLOEMLKMFQMSGPEKLALDELIDSCPTQVKH 189
QY 132 INSYLKPMLQDRFETALPARGLDIAENILSYLDKSLCAAEVLCYKEMRYVTSQGMKMK 191
Db 190 MMQVTEPQFQDFISLIP-----KELALYVLSFLEPKDLQAQOTCRWRIIAEDNLMRE 245
QY 192 LIEMWRTSLMRGLAE-----RRCMGQYLFKNRPPDGNAPPNSTYRALLYPKIIODIETI 246
Db 246 KCKE-----EGIDEPHLIKRRK-----VTKPGFIHSPKMSAY-----IRO--HRI 283
QY 247 ESNMRCGRHSIORICRSETSKGYCLOYDDOKIYSGLRDNTIKIMDKNTLECKRIILGH 306
Db 284 DNNMRGELKSPKV-LKGHDHVITCLOFGNRIYSGSDNTLKYMSAVTGKCLRTLVGH 342
QY 307 TGSVLCLOYDERVIITSSSDSTVRYWVDNTGEMTLIIHCEAVLHLRFNNGMMVYCSKD 366
Db 343 TGGVWSSQMRDNIISGSTDTLTKVWNAETGECIHTLYGHTSVRCMHLHEKRYVSGSRD 402
QY 367 RSLAVWMDASPTDITLRVYLGHRRAAVNVDPDDKXIYASGDRITIKVWNTSTCEPVRTL 426
Db 403 ATLWMDIETGQCL--HYLMGHVAAVRCVQYDGRVVSAGVDFMKVWDEPTEETCLHTL 459
QY 427 NGHKRGIACTQYRDLRYVSSSDNTIRLMDIEGACILRVLEGEHELIVRCIRFNDKRIYSG 486
Db 460 QGHTNRVYSLQFDGIHVYSGSLDTSIRWVDVETGNCIHTLGHQSLTSGMELKONILVSG 519
QY 487 AYDGKIKVMDVLAALDPRAAGTLCRLTY---EHSGRVFRLOPDEFQIVSSSHDDTILI 543
Db 520 NADSTVKIMDIKTG-----QCLQTLQGPKNHOSAVTCLQFNKNFVITSSDDGTIVKL 570
QY 544 WD 545
Db 571 WD 572

```

Search completed: February 20, 2003, 09:55:19

Job time : 44 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:58:07 : Search time 14 Seconds
(without alignments)
1038.379 Million cell updates/sec

Title: US-09-601-168b-2
Perfect score: 3034
Sequence: 1 MDPAAVQEKALKFMSNSE.....PAAQAEPPSPSRITYISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications -AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3034	100.0	569	12	US-10-042-417-2
2	1163	38.3	219	9	US-10-023-530-2
3	989.5	32.6	265	10	US-09-764-848-30
4	644.5	21.2	626	9	US-09-213-888-21
5	644.5	21.2	626	9	US-09-328-877A-21
6	640	21.1	540	9	US-09-213-888-7
7	640	21.1	540	9	US-09-213-888-10
8	640	21.1	540	9	US-09-328-877A-7
9	640	21.1	540	9	US-09-328-877A-10
10	640	21.1	545	9	US-09-213-888-6
11	640	21.1	545	9	US-09-328-877A-6
12	640	21.1	553	9	US-09-213-888-5
13	640	21.1	553	9	US-09-328-877A-5
14	640	21.1	559	9	US-09-213-888-9
15	640	21.1	559	9	US-09-328-877A-9
16	640	21.1	589	9	US-09-213-888-8
17	640	21.1	589	9	US-09-328-877A-8
18	640	21.1	592	9	US-09-213-888-4
19	640	21.1	592	9	US-09-328-877A-4

20	640	21.1	627	9	US-09-213-888-3	Sequence 3, Appl1
21	640	21.1	627	9	US-09-328-877A-3	Sequence 2, Appl1
22	640	21.1	666	9	US-09-213-888-27	Sequence 27, Appl1
23	640	21.1	666	9	US-09-328-877A-27	Sequence 27, Appl1
24	640	21.1	669	9	US-09-213-888-25	Sequence 25, Appl1
25	640	21.1	669	9	US-09-328-877A-25	Sequence 25, Appl1
26	575	19.0	678	10	US-09-801-368-314	Sequence 314, App
27	545	18.0	678	10	US-10-060-019-30	Sequence 30, Appl1
28	455.5	15.0	1356	9	US-10-077-111-10	Sequence 10, Appl1
29	399	13.2	779	9	US-10-060-019-29	Sequence 29, Appl1
30	392	12.9	732	10	US-09-994-485-8	Sequence 8, Appl1
31	384	12.7	732	9	US-09-832-292-12	Sequence 12, Appl1
32	337	11.1	1146	9	US-09-994-485-6	Sequence 6, Appl1
33	337	11.1	1146	10	US-10-042-417-4	Sequence 4, Appl1
34	334.5	11.0	422	12	US-10-132-744A-4	Sequence 6, Appl1
35	316.5	10.4	485	9	US-10-132-744A-6	Sequence 6, Appl1
36	307.5	10.1	742	9	US-10-077-111-11	Sequence 11, Appl1
37	299.5	9.9	1194	10	US-09-876-667-2	Sequence 2, Appl1
38	299.5	9.9	1205	10	US-09-876-667-16	Sequence 16, Appl1
39	296.5	9.8	484	9	US-10-132-744A-2	Sequence 2, Appl1
40	293.5	9.7	261	9	US-10-132-744A-4	Sequence 4, Appl1
41	283	9.3	713	10	US-09-801-368-408	Sequence 408, App
42	264.5	8.7	423	10	US-09-729-674-160	Sequence 160, App
43	260.5	8.6	375	9	US-10-119-932-1	Sequence 1, Appl1
44	254.5	8.4	316	10	US-09-828-310-12	Sequence 12, Appl1
45	246.5	8.1	343	9	US-10-119-932-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-10-042-417-2
Sequence 2, Application US/10042417
Patent No. US2002123082A1
GENERAL INFORMATION:
APPLICANT: Patano, M.
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
FILE REFERENCE: 5914-090-999
CURRENT APPLICATION NUMBER: US/10/042,417
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: 60/260,179
PRIOR FILING DATE: 2001-01-5
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 569
TYPE: PRT
ORGANISM: Homo sapiens
US-10-042-417-2

Query Match 100.0%; Score 3034; DB 12; Length 569;
Best Local Similarity 100.0%; Pred. No. 6.5e-248;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDPAAVQEKALKFMSNSEPPKRTIPKNSLRQFYNSCARCLNOETVCLA	60
DB	1	MDPAAVQEKALKFMSNSEPPKRTIPKNSLRQFYNSCARCLNOETVCLA	60
QY	61	STAMKTEVCATKTLANGTSSMIVPQKRLSASYEKEKELCVYFEQWSSDQVEVEHL	120
DB	61	STAMKTEVCATKTLANGTSSMIVPQKRLSASYEKEKELCVYFEQWSSDQVEVEHL	120
QY	121	ISOMCHYGHGHTNSYLKPMLODFTTALPANGLDHIAENILSYLDASLSCAAEVCKEWY	180
DB	121	ISOMCHYGHGHTNSYLKPMLODFTTALPANGLDHIAENILSYLDASLSCAAEVCKEWY	180
QY	181	RYTSOSGLMKKLTIEHVTRDLSMRGLAERRGQYLFNKKPPDGNAPNSFYRALYKII	240
DB	181	RYTSOSGLMKKLTIEHVTRDLSMRGLAERRGQYLFNKKPPDGNAPNSFYRALYKII	240
QY	241	ODIETIESNMCRGRSLORHCRSETSKGVYCLQYDQKIVSGLRDNTIKITWKDNTLECK	300

```

Db 241 QDIETESNMRCGRHSLSORICRSETSGVYCIQYDQKIVSGRDNITIKIMKNTLECK 300
QY 301 RILHTGHSVLCLOYDEVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVLHLPFNNGMM 360
Db 301 RILHTGHSVLCLOYDEVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVLHLPFNNGMM 360
QY 361 VTCSKRSIAVWMAFPTDITLRVLVGHRAAVVDFEDKTYVASGDDTIVWNTSTG 420
Db 361 VTCSKRSIAVWMAFPTDITLRVLVGHRAAVVDFEDKTYVASGDDTIVWNTSTG 420
QY 421 EFVTLNKHKGICLQYRDRLVYSSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
Db 421 EFVTLNKHKGICLQYRDRLVYSSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIISGAVDGKIKYWDIVLADLPAPAGTICLRVLVHSGRVRLQDFEQLVSSSHDT 540
Db 481 KRIISGAVDGKIKYWDIVLADLPAPAGTICLRVLVHSGRVRLQDFEQLVSSSHDT 540
QY 541 ILIMDFLNDPAQAEPSPSRRTTYISR 569
Db 541 ILIMDFLNDPAQAEPSPSRRTTYISR 569

```

RESULT 2

```

US-10-023-530-2
; Sequence 2, Application US/10023530
; Publication No. US20030007956A1
; GENERAL INFORMATION:
; APPLICANT: LEGRAIN, Pierre
; APPLICANT: BENAROUS, Richard
; APPLICANT: BLOT, Guillaume
; APPLICANT: LASSOT, Irina
; TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP
; FILE REFERENCE: B4717A
; CURRENT APPLICATION NUMBER: US/10/023,530
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 60/256,276
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: BetaTrCP
; LOCATION: (1)..(219)
; OTHER INFORMATION: F-box protein
US-10-023-530-2

```

```

Query Match          38.3%; Score 1163; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 9, 6e-91;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLQERKALFKMNSEREDCNNGEPKRIIPKNSLRQYNSCARLCLNOETVCLA 60
Db 1 MDPAAVLQERKALFKMNSEREDCNNGEPKRIIPKNSLRQYNSCARLCLNOETVCLA 60
QY 61 STAMTEKCVAKTKLANSTSMIYVKKOKLSASYEKELCVKFFEQWSESDVEVEEHL 120
Db 61 STAMTEKCVAKTKLANSTSMIYVKKOKLSASYEKELCVKFFEQWSESDVEVEEHL 120
QY 121 ISOMCHYOHGHSINSLKPMLODRFTLPAKGLDHAENILSLYDAKSLCAAEVCKEWM 180
Db 121 ISOMCHYOHGHSINSLKPMLODRFTLPAKGLDHAENILSLYDAKSLCAAEVCKEWM 180
QY 181 RYTSQGMIMKKLIERWVRTDSLWMLGAEKRGWQYLFKN 219
Db 181 RYTSQGMIMKKLIERWVRTDSLWMLGAEKRGWQYLFKN 219

```

RESULT 3

```

US-09-764-848-30
; Sequence 30, Application US/09764848
; Patent No. US20020077270A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT208
; CURRENT APPLICATION NUMBER: US/09/764,848
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 265
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-764-848-30

```

Query Match

```

Best Local Similarity 32.6%; Score 989.5; DB 10; Length 265;
Matches 186; Conservative 24; Mismatches 21; Indels 39; Gaps 3;

```

```

QY 15 FMNSEREDCNNGEPKRIIPKNSLRQYNSCARLCLNOETVCLASTAMKTEKVAKTK 74
Db 20 FQNTSVMEQDNEDSPK-----KNTLMO-----42
QY 75 LANGTSMIYVKKOKLSASYEKELCVKFFEQWSESDVEVEEHLISOMCHYOHGHSIN 134
Db 43 ISNGTSVIVSRKRPRSEGNYQKEKDCIKYFPQWSESDVEVEEHLISOMCHYOHGHSIN 102
QY 135 YLKPMLQDRFTLPAKGLDHAENILSLYDAKSLCAAEVCKEWMRYTSDGMLMKLIE 194
Db 103 YLKPMLQDRFTLPAKGLDHAENILSLYDAKSLCAAEVCKEWMRYTSDGMLMKLIE 162
QY 195 RMYRTDSLWMLGAEKRGWQYLFKNKPPGNAPNSFYRALYPTIODETIESNMRCGR 254
Db 163 RMYRTDSLWMLGAEKRGWQYLFKNKPPGNAPNSFYRALYPTIODETIESNMRCGR 220
QY 255 HSLQRIHCRSEKSGVYCIQYDQKIVSG 284
Db 221 HSLQRIHCRSEKSGVYCIQYDQKIVSG 250

```

RESULT 4

```

US-09-213-888-21
; Sequence 21, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 626
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6 myc tagged
; OTHER INFORMATION: homo sapien
US-09-213-888-21

```

Query Match

```

Best Local Similarity 21.2%; Score 644.5; DB 9; Length 626;
Matches 169; Conservative 103; Mismatches 222; Indels 67; Gaps 15;

```

```

QY 8 LQERKALFKMNSEREDCNNGEPKRIIPK--NSLRQ--TYNSCARLCLNOETVCLASTA 63

```



```

Db 101 MMQVLEPQFQDFISLIP-----KELALYVLSFLEPKDILLQAQTCRYWRIIAEDNLLMRE 156
QY 192 LIERNVTRDISLWRGLAE-----RKGWGYLFKNRPPDGNAPPNSFYALYPKIIODIETI 246
Db 157 KCKE-----EGIDPELHTRKR-----VIKPGFIHSPWKSAY-----IRQ--HRI 194
QY 247 ESNMRCGRHSIORIHCRSETSKGYCLOQYDDOKIVSGLRDNTIKIMKNTLECKRIITLGH 306
Db 195 DTNMRGELKSPKV-LKQHDHVTICLOFCGNRIYVSGSDNTILKWSAAYGKCLRTILVGH 253
QY 307 TGSVLCLOYDERVITITGSSDSTVAVDNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366
Db 254 TGGVWSSQMRDNIITIGSDTRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRYVSGSRD 313
QY 367 RSIAYWDMASPTDITLRVLYGHRAAVNVDPDDKYIVASGDRITKWNNTSTCEFRVTL 426
Db 314 ATLWMDIEFGQCL---HYLMGHVAAVRCVOYDGRVYSGAYDPMVWMDPETETCLHTL 370
QY 427 NGHRGIACTLOYRDLRYVSGSSDNTIRLMDIEGACLRLEHEELVRCIRFDNKRIVSG 486
Db 371 OGHTRNVYSILOPDGIVHVSGLDTSIRWMDVETGNCIHTLGHQSLTSGMELKDNILVSG 430
QY 487 AYDGRKIVMDVLAALDPRAPAGTLCRTLV---EHSGRVRLQDFEQIVSSSHDITLI 543
Db 431 NADSTVKIMDIKTG-----QCLQTLQGNPKHQSAYVTCLOFNKNFVITSSDDGTIVKL 481
QY 544 WD 545
Db 482 WD 483

```

```

RESULT 7
US-09-213-888-10
; Sequence 10, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-10

```

```

Query Match 21.1%; Score 640; DB 9; Length 540;
Best Local Similarity 30.8%; Pred. No. 4.1e-46;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```

```

QY 14 KFMNSREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQEIVCLASTAKTENCYAKT 73
Db 4 KLDHSEVRSFSLGKKPKV-----SEYTSITGL-----VPCSA-----TPPTFGDL 45
QY 74 KLANGTSSMIVPKOKRLSASYEKEL--CYKFEQWSESDQVEFEVHLISQMCYOHGH 131
Db 46 RAANGOG-----QQRRTITSVOPPTGIQEWLKMFSWGSPEKLLADELIDSCPTQYKH 100
QY 132 INSYLKPMLQDFITALPARGLDHAENILSYLDAKSLCAELVCKEYRVTSDGMLMK 191
Db 101 MMQVLEPQFQDFISLIP-----KELALYVLSFLEPKDILLQAQTCRYWRIIAEDNLLMRE 156
QY 192 LIERNVTRDISLWRGLAE-----RKGWGYLFKNRPPDGNAPPNSFYALYPKIIODIETI 246
Db 157 KCKE-----EGIDPELHTRKR-----VIKPGFIHSPWKSAY-----IRQ--HRI 194

```

```

QY 247 ESNMRCGRHSIORIHCRSETSKGYCLOQYDDOKIVSGLRDNTIKIMKNTLECKRIITLGH 306
Db 195 DTNMRGELKSPKV-LKQHDHVTICLOFCGNRIYVSGSDNTILKWSAAYGKCLRTILVGH 253
QY 307 TGSVLCLOYDERVITITGSSDSTVAVDNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366
Db 254 TGGVWSSQMRDNIITIGSDTRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRYVSGSRD 313
QY 367 RSIAYWDMASPTDITLRVLYGHRAAVNVDPDDKYIVASGDRITKWNNTSTCEFRVTL 426
Db 314 ATLWMDIEFGQCL---HYLMGHVAAVRCVOYDGRVYSGAYDPMVWMDPETETCLHTL 370
QY 427 NGHRGIACTLOYRDLRYVSGSSDNTIRLMDIEGACLRLEHEELVRCIRFDNKRIVSG 486
Db 371 OGHTRNVYSILOPDGIVHVSGLDTSIRWMDVETGNCIHTLGHQSLTSGMELKDNILVSG 430
QY 487 AYDGRKIVMDVLAALDPRAPAGTLCRTLV---EHSGRVRLQDFEQIVSSSHDITLI 543
Db 431 NADSTVKIMDIKTG-----QCLQTLQGNPKHQSAYVTCLOFNKNFVITSSDDGTIVKL 481
QY 544 WD 545
Db 482 WD 483

```

```

RESULT 8
US-09-328-877A-7
; Sequence 7, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-7

```

```

Query Match 21.1%; Score 640; DB 9; Length 540;
Best Local Similarity 30.8%; Pred. No. 4.1e-46;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```

```

QY 14 KFMNSREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQEIVCLASTAKTENCYAKT 73
Db 4 KLDHSEVRSFSLGKKPKV-----SEYTSITGL-----VPCSA-----TPPTFGDL 45
QY 74 KLANGTSSMIVPKOKRLSASYEKEL--CYKFEQWSESDQVEFEVHLISQMCYOHGH 131
Db 46 RAANGOG-----QQRRTITSVOPPTGIQEWLKMFSWGSPEKLLADELIDSCPTQYKH 100
QY 132 INSYLKPMLQDFITALPARGLDHAENILSYLDAKSLCAELVCKEYRVTSDGMLMK 191
Db 101 MMQVLEPQFQDFISLIP-----KELALYVLSFLEPKDILLQAQTCRYWRIIAEDNLLMRE 156
QY 192 LIERNVTRDISLWRGLAE-----RKGWGYLFKNRPPDGNAPPNSFYALYPKIIODIETI 246
Db 157 KCKE-----EGIDPELHTRKR-----VIKPGFIHSPWKSAY-----IRQ--HRI 194
QY 247 ESNMRCGRHSIORIHCRSETSKGYCLOQYDDOKIVSGLRDNTIKIMKNTLECKRIITLGH 306
Db 195 DTNMRGELKSPKV-LKQHDHVTICLOFCGNRIYVSGSDNTILKWSAAYGKCLRTILVGH 253
QY 307 TGSVLCLOYDERVITITGSSDSTVAVDNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366

```


Db 254 TGVWSSQMRDNIISGSTRDLKLVMAETGECIHTLYGHTYRVCMLHKEKRVSSRD 313
 QY 367 RSLAVWDMASPDITLRLVGLVHRAAVVVDDEDDIYASAGDRTIKWNTSTCEFRTL 426
 Db 314 ATLWMDIETGCL---HVLKMGHVAARVCGYDGRVVSAGAYDFWVKWMDPETETCLHTL 370
 QY 427 NGHKGACIQLYRDLRVSSGSDNTIRLMDIEGACLVLEBGEHLVLCIFPDNKRIVSG 486
 Db 371 OGHTRNVSLSQFDGIHVYSGSLDTSIRWMDVETGNCIHTLGHOSLISGMELKDNILVSG 430
 QY 487 AYDGKIKWMDVLAALDPRAPAGTLCRLTV---EHSGRVFLQDFEFOYSSSHDDTILI 543
 Db 431 NADSTVAKIMDIKTG-----OCLQTLQGNKQSAVTCLOFKNKVNITSSDDGTIVKL 481
 QY 544 WD 545
 Db 482 WD 483

RESULT 9
 US-09-328-877A-10
 : Sequence 10, Application US/09328877A
 : Patent No. US20020177187A1
 : GENERAL INFORMATION:
 : APPLICANT: Gurney, Mark E.
 : APPLICANT: Li, Jinhe
 : APPLICANT: Pauley, Adele M.
 : APPLICANT: Pharmacia & Upjohn Company
 : TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
 : FILE REFERENCE: 6142
 : CURRENT APPLICATION NUMBER: US/09/328,877A
 : CURRENT FILING DATE: 1999-06-09
 : NUMBER OF SEQ ID NOS: 27
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 10
 : LENGTH: 540
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-09-328-877A-10

Query Match 21.1%; Score 640; DB 9; Length 540;
 Best Local Similarity 30.8%; Pred. No. 4.1e-46;
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KFMNSEREDCNNGEPPKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTEHCYAKT 73
 Db 4 KLDHGEVRSFSLKPKCKV-----SEYTSITGL-----VPCSA-----TPPTFGDL 45
 QY 74 KLANGTSSMIVPKOKRLSASYEKEKEL--CYKFEQMSSESDOVEFEVHLISOMCHVOGH 131
 Db 46 RAANGOG-----QORRITSVQPTGLQEWLKMFOMSGPEKLLALDELIDSCPTQVKH 100
 QY 132 INSYLKPMLQDRFTALPARGLDIAENILSYLDAKSICAAELVCKEMRYVTSGLMLWK 191
 Db 101 MMQVIEPQRFQDFISLP---KELALYVLSFLEPKDLQAAQTCRWYRIILAEENLIMRE 156
 QY 192 LIERVATRDSLMKGLAE-----RGMGOYLFPKNKPPDGNAPNSFYALYPKIIODIETI 246
 Db 157 KCKE-----EGIDELPLHKRRK-----VIRGFHSHWKSAY-----IRQ--HRI 194
 QY 247 ESNMRCGRHSIORHICRSETSKGVYCLQYDDOKIYVGLRDNITIKIMDKNLECKRILITGH 306
 Db 195 DTNMRGELKSPKV-LKGDHDVITICLOFCGNRIYSGSDDTLTKWMSAVTGKCLRTLVGH 253
 QY 307 TGSVLCLOQDERVITITGSSDSTVRWMDVNTGEMLTLIHCEAVLHLRFNNGMMVYTCSD 366
 Db 254 TGVWSSQMRDNIISGSTRDLKLVMAETGECIHTLYGHTYRVCMLHKEKRVSSRD 313
 QY 367 RSLAVWDMASPDITLRLVGLVHRAAVVVDDEDDIYASAGDRTIKWNTSTCEFRTL 426
 Db 314 ATLWMDIETGCL---HVLKMGHVAARVCGYDGRVVSAGAYDFWVKWMDPETETCLHTL 370

QY 427 NGHKGACIQLYRDLRVSSGSDNTIRLMDIEGACLVLEBGEHLVLCIFPDNKRIVSG 486
 Db 371 OGHTRNVSLSQFDGIHVYSGSLDTSIRWMDVETGNCIHTLGHOSLISGMELKDNILVSG 430
 QY 487 AYDGKIKWMDVLAALDPRAPAGTLCRLTV---EHSGRVFLQDFEFOYSSSHDDTILI 543
 Db 431 NADSTVAKIMDIKTG-----OCLQTLQGNKQSAVTCLOFKNKVNITSSDDGTIVKL 481
 QY 544 WD 545
 Db 482 WD 483

RESULT 10
 US-09-213-888-6
 : Sequence 6, Application US/09213888A
 : Patent No. US20020164683A1
 : GENERAL INFORMATION:
 : APPLICANT: Gurney, Mark E.
 : APPLICANT: Li, Jinhe
 : APPLICANT: Pauley, Adele M.
 : APPLICANT: Pharmacia & Upjohn Company
 : TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
 : FILE REFERENCE: 6142
 : CURRENT APPLICATION NUMBER: US/09/213,888A
 : CURRENT FILING DATE: 1998-12-17
 : NUMBER OF SEQ ID NOS: 27
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 6
 : LENGTH: 545
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-09-213-888-6

Query Match 21.1%; Score 640; DB 9; Length 545;
 Best Local Similarity 30.8%; Pred. No. 4.2e-46;
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KFMNSEREDCNNGEPPKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTEHCYAKT 73
 Db 9 KLDHGEVRSFSLKPKCKV-----SEYTSITGL-----VPCSA-----TPPTFGDL 50
 QY 74 KLANGTSSMIVPKOKRLSASYEKEKEL--CYKFEQMSSESDOVEFEVHLISOMCHVOGH 131
 Db 51 RAANGOG-----QORRITSVQPTGLQEWLKMFOMSGPEKLLALDELIDSCPTQVKH 105
 QY 132 INSYLKPMLQDRFTALPARGLDIAENILSYLDAKSICAAELVCKEMRYVTSGLMLWK 191
 Db 106 MMQVIEPQRFQDFISLP---KELALYVLSFLEPKDLQAAQTCRWYRIILAEENLIMRE 161
 QY 192 LIERVATRDSLMKGLAE-----RGMGOYLFPKNKPPDGNAPNSFYALYPKIIODIETI 246
 Db 162 KCKE-----EGIDELPLHKRRK-----VIRGFHSHWKSAY-----IRQ--HRI 199
 QY 247 ESNMRCGRHSIORHICRSETSKGVYCLQYDDOKIYVGLRDNITIKIMDKNLECKRILITGH 306
 Db 200 DTNMRGELKSPKV-LKGDHDVITICLOFCGNRIYSGSDDTLTKWMSAVTGKCLRTLVGH 258
 QY 307 TGSVLCLOQDERVITITGSSDSTVRWMDVNTGEMLTLIHCEAVLHLRFNNGMMVYTCSD 366
 Db 259 TGVWSSQMRDNIISGSTRDLKLVMAETGECIHTLYGHTYRVCMLHKEKRVSSRD 318
 QY 367 RSLAVWDMASPDITLRLVGLVHRAAVVVDDEDDIYASAGDRTIKWNTSTCEFRTL 426
 Db 319 ATLWMDIETGCL---HVLKMGHVAARVCGYDGRVVSAGAYDFWVKWMDPETETCLHTL 375
 QY 427 NGHKGACIQLYRDLRVSSGSDNTIRLMDIEGACLVLEBGEHLVLCIFPDNKRIVSG 486
 Db 376 OGHTRNVSLSQFDGIHVYSGSLDTSIRWMDVETGNCIHTLGHOSLISGMELKDNILVSG 435
 QY 487 AYDGKIKWMDVLAALDPRAPAGTLCRLTV---EHSGRVFLQDFEFOYSSSHDDTILI 543

Db 436 NADSTVKIMDIKTG-----QCLQTLQGNPKHOSAVTCLQFNKNFVITSSDGTIVKL 486
QY 544 WD 545
Db 487 WD 488

RESULT 11
US-09-328-877A-6
; Sequence 6, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142 Encode Them
; CURRENT APPLICATION NUMBER: US/09/328, 877A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-6

Query Match 21.1%; Score 640; DB 9; Length 545;
Best Local Similarity 30.8%; Pred. No. 4.2e-46;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KMNSSEREDCNCEPPKRIPEKNSLRQYNSCARLQNEVYCLASTAMKTEVCYAKT 73
Db 9 KLDHSEVRSSFLGKKPKV-----SEYTTSTGL-----VPCSA-----TPTTFGDL 50
QY 74 KLANGTSSAIYPKOKKLSASTYKEKEL--CYKFEQMSSEDOVEFEVHLISQCHQOHH 131
Db 51 RAANGOG-----QQRRTTSVQPTGLQEWLKMFOGWSGPEKLLAIDELIDSEPEPOVKA 105
QY 132 INSYLKPMLQDFETALPARGLDHIAENILSYLDAKSLCAELVCKEYKRVSDGMLMK 191
Db 106 MMQVTEPQFQDFISLTP-----KELALYVLSFEPRKDLQAOTCRWRILAEIDLME 161
QY 192 LIERNVRTDSLWRGLAE-----RRWGQOYLFPKKNPPDGNAPPNPFRALYPKTIIDIEFI 246
Db 162 KCKE-----EGIDPEPLHKRRK-----VIKPGFIHSPKMSAV-----IRO--HRI 199
QY 247 ESNMRCGRHSIORHCRSETSKGYVCLQYDQKIVSGLRDNTIKIDKNTLECKRILTGH 306
Db 200 DTNMRGELKSPKV-LKGHDDHVTCLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLVGH 258
QY 307 TGSVLCLOYDERVITIGSSDSYRVMDVNTGEMLNTLIHHCNAVHLIRFNNGMATYCSKD 366
Db 259 TGGWSSQMRDNIITISGSTDRITLKVNAETGECIHTLYGHTSTVRCMHLHEKRVYSGSD 318
QY 367 RSLAVMDASPTDITLRLVYGHRAAVNVPDQKIYASGDRITLKVNTSTCEVFTL 426
Db 319 ATLRYMDIEIGQCL--HYLMGHVAAVRCVQYDGRVYSGADPFWKWDPEETETCLHTL 375
QY 427 NGHKGACIQLQYDRILVYSGSSDNTIRLMDIEGACRLVLEHHELVYRCIRDNKRIYSG 486
Db 376 QGHTNRYYSIQDFGIHVYSGSLDTSIRVMDVETGNCIHTLYGHQSLTSGMELKMDIIVSG 435
QY 487 AYDGKIKVMDLVAALDPRAPAGTLCRLTV---EHSGRFRLQDFEFOIVSSSHDITLI 543
Db 436 NADSTVKIMDIKTG-----QCLQTLQGNPKHOSAVTCLQFNKNFVITSSDGTIVKL 486
QY 544 WD 545
Db 487 WD 488

RESULT 12
US-09-213-888-5
; Sequence 5, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142 Encode Them
; CURRENT APPLICATION NUMBER: US/09/213, 888A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-5

Query Match 21.1%; Score 640; DB 9; Length 553;
Best Local Similarity 30.8%; Pred. No. 4.3e-46;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KMNSSEREDCNCEPPKRIPEKNSLRQYNSCARLQNEVYCLASTAMKTEVCYAKT 73
Db 17 KLDHSEVRSSFLGKKPKV-----SEYTTSTGL-----VPCSA-----TPTTFGDL 58
QY 74 KLANGTSSAIYPKOKKLSASTYKEKEL--CYKFEQMSSEDOVEFEVHLISQCHQOHH 131
Db 59 RAANGOG-----QQRRTTSVQPTGLQEWLKMFOGWSGPEKLLAIDELIDSEPEPOVKA 113
QY 132 INSYLKPMLQDFETALPARGLDHIAENILSYLDAKSLCAELVCKEYKRVSDGMLMK 191
Db 114 MMQVTEPQFQDFISLTP-----KELALYVLSFEPRKDLQAOTCRWRILAEIDLME 169
QY 192 LIERNVRTDSLWRGLAE-----RRWGQOYLFPKKNPPDGNAPPNPFRALYPKTIIDIEFI 246
Db 170 KCKE-----EGIDPEPLHKRRK-----VIKPGFIHSPKMSAV-----IRO--HRI 207
QY 247 ESNMRCGRHSIORHCRSETSKGYVCLQYDQKIVSGLRDNTIKIDKNTLECKRILTGH 306
Db 208 DTNMRGELKSPKV-LKGHDDHVTCLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLVGH 266
QY 307 TGSVLCLOYDERVITIGSSDSYRVMDVNTGEMLNTLIHHCNAVHLIRFNNGMATYCSKD 366
Db 267 TGGWSSQMRDNIITISGSTDRITLKVNAETGECIHTLYGHTSTVRCMHLHEKRVYSGSD 326
QY 367 RSLAVMDASPTDITLRLVYGHRAAVNVPDQKIYASGDRITLKVNTSTCEVFTL 426
Db 327 ATLRYMDIEIGQCL--HYLMGHVAAVRCVQYDGRVYSGADPFWKWDPEETETCLHTL 383
QY 427 NGHKGACIQLQYDRILVYSGSSDNTIRLMDIEGACRLVLEHHELVYRCIRDNKRIYSG 486
Db 384 QGHTNRYYSIQDFGIHVYSGSLDTSIRVMDVETGNCIHTLYGHQSLTSGMELKMDIIVSG 443
QY 487 AYDGKIKVMDLVAALDPRAPAGTLCRLTV---EHSGRFRLQDFEFOIVSSSHDITLI 543
Db 444 NADSTVKIMDIKTG-----QCLQTLQGNPKHOSAVTCLQFNKNFVITSSDGTIVKL 494
QY 544 WD 545
Db 495 WD 496

RESULT 13
US-09-328-877A-5
; Sequence 5, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:


```

; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-9

```

Query Match	21.18;	Score 640;	DB 9;	Length 559;
-------------	--------	------------	-------	-------------

```

QY 14 KFMNSSEEDCNDNEPPK11PEKNRS1RQYNSCARCLNDQFVCLASTAMKTENCVAKT 73
Db 23 KLDHGEVRSRSLCKRCKV-----SETSTTGL-----VPCSA-----1PTTGD 64
QY 74 KLANGTSSIMVPRKJLSASYEKKEK1--CVKYEPQMSSEDOVEFEHLLSOMCHYOGH 131
Db 65 RAAMQOC-----QQRRTTSVQPTGLOEWLKMFOQMSGPKRLALDELIDSEPTGVKH 119
QY 132 INSTLKPMLQDFTTALPARGLDIAENILSYLDAKSLCAELVCKEMYRYTSDGMIMK 191
Db 120 MMOVIEPQFQDFTSLP1--KELALVYSLFERKDLLQAQTCRYWRLIAEDNLLMRE 179
QY 192 LIERNVRDTSLMRGLAE-----RQMGQYLFKNRPDQNAAPPNSFYBALPKIIDLETI 246
Db 176 KCKE-----EGIDPELHKRRK-----VYKPGFHSBWSAY-----IRQ-HRI 213
QY 247 ESNRRCGRHS1OR1HCRES7SKGYCLOYDQKIVSGLRNDTIKIMDKNLTLECKRILTGH 306
Db 214 DTNNRGRGLSPKV-LKGDHDHVITCLOFGCNRIYVSSDDNTLKWASAAYGKCLRTLYGH 272
QY 307 TGSVLCLOYDERVYITGSSDSTVYAVDVTNGEMTL1IHCEAVLHLRFNNGMKVYTSKD 366
Db 273 TGGVWSSQMRNIIISGSTDRTLLKVMNAEGEC1HTLYGHTSTVRCMHLHEKRVSSGRD 332
QY 367 RSIWVMDASPTDITLRLVYLGHAAVNVVDQDKYIVASQDPTIKVMN1SCEPRTL 426
Db 333 ATLNRWDETQOCL--HVLMGHAAARCVQYDGRRVASGYADYDVXWMPDETETCLHTL 389
QY 427 NGHRKGIACLOYRDLVWSSGSDTIRLMDIEGCACLRVLEGHEELRCAIRFDNRKRVSG 486
Db 390 QGH1NRRVYSLOFQDGLIHVYSSGLD1S1KVMYDEITGNC1HTLHGSLTSGMELKDNLLVSG 449
QY 487 AYDKIKVMDVLAALDPRAAGTLCRLTV1--EHSGRVRLQDDEFQIVSSSHDITLT 543
Db 450 NADSTVXIMDKTG-----QCLQTLQPNKHQSAVTCLOFNKNFVLTSSDDGTVKL 500
QY 544 WD 545
Db 501 WD 502

```

Search completed: February 20, 2003, 10:00:10
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:56:12 : Search time 57 Seconds

(without alignments)
897.640 Million cell updates/sec

Title: US-09-601-168B-2

Perfect score: 3034

Sequence: 1 MDPRAVLOEKALKFNMSSEREDCNNGEPKRIIPKNSLRQYNSCARLCAQETVCA 569

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 468380 seqs, 89921735 residues

Total number of hits satisfying chosen parameters: 468380

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score of the result being printed, and is derived by analysis of the score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3034	100.0	569	US-09-601-168B-2	Sequence 2, Appl
2	3006	99.1	605	US-09-724-676-52731	Sequence 52731, A
3	3006	99.1	605	US-09-724-676-52731	Sequence 52731, A
4	2701	89.0	537	US-09-724-676-52729	Sequence 52729, A
5	2701	89.0	537	US-09-724-676-52730	Sequence 52730, A
6	2701	89.0	537	US-09-724-676-52729	Sequence 52729, A
7	2701	89.0	537	US-09-724-676-52730	Sequence 52730, A
8	2673	88.1	573	US-09-724-676-52732	Sequence 52732, A
9	2673	88.1	573	US-09-724-676-52732	Sequence 52732, A
10	2673	88.1	573	US-09-724-676-52733	Sequence 52733, A
11	2673	88.1	573	US-09-724-676-52733	Sequence 52733, A
12	2400	79.1	529	US-09-724-676-61015	Sequence 61015, A
13	2400	79.1	529	US-09-724-676-61016	Sequence 61016, A
14	2400	79.1	529	US-09-724-676-61015	Sequence 61015, A
15	2400	79.1	529	US-09-724-676-61016	Sequence 61016, A
16	2384.5	78.6	542	US-09-724-676-61014	Sequence 61014, A
17	2384.5	78.6	542	US-09-724-676-61014	Sequence 61014, A
18	2358	77.7	512	US-09-724-676-61018	Sequence 61018, A
19	2358	77.7	512	US-09-724-676-61018	Sequence 61018, A
20	2358	77.7	512	US-09-724-676-61017	Sequence 61017, A
21	2358	77.7	512	US-09-724-676-61018	Sequence 61018, A
22	545	18.0	640	US-10-060-019-30	Sequence 30, Appl
23	399	13.2	779	US-10-060-019-29	Sequence 29, Appl
24	354	11.7	410	US-09-724-676-61377	Sequence 61377, A
25	354	11.7	410	US-09-724-676-61378	Sequence 61378, A
26	354	11.7	410	US-09-724-676-61379	Sequence 61379, A

27	354	11.7	410	US-09-724-676A-61377	Sequence 61377, A
28	354	11.7	410	US-09-724-676A-61378	Sequence 61378, A
29	354	11.7	410	US-09-724-676A-61379	Sequence 61379, A
30	334.5	11.0	486	US-09-724-676-60984	Sequence 60984, A
31	334.5	11.0	486	US-09-724-676A-60984	Sequence 60984, A
32	334.5	11.0	600	US-10-197-666A-120	Sequence 120, App
33	334.5	11.0	631	US-10-197-666A-116	Sequence 116, App
34	334.5	11.0	667	US-10-197-666A-118	Sequence 118, App
35	334.5	11.0	670	US-10-197-666A-114	Sequence 114, App
36	332.5	11.0	357	US-10-264-237-2768	Sequence 2768, Ap
37	330.5	10.9	631	US-10-197-666A-144	Sequence 144, App
38	330.5	10.9	641	US-10-197-666A-112	Sequence 112, App
39	330.5	10.9	670	US-10-197-666A-150	Sequence 150, App
40	330.5	10.9	670	US-10-218-140-4360	Sequence 4360, Ap
41	324.5	10.7	334	PCT-US02-27671A-2	Sequence 2, Appl
42	324.5	10.7	334	PCT-US02-27671A-4	Sequence 4, Appl
43	314	10.3	524	PCT-US02-40925-3221	Sequence 3221, Ap
44	314	10.3	524	US-10-320-797-3221	Sequence 3221, Ap
45	303.5	10.0	1213	US-09-724-676-50571	Sequence 50571, A

ALIGNMENTS

RESULT 1
US-09-601-168B-2
Sequence 2, Application US/09601168B
GENERAL INFORMATION:
APPLICANT: BENAROUS, Richard
APPLICANT: MARGOTIN, Florence
APPLICANT: DURAND, Hervé
APPLICANT: AREZVANA SEISDEDOS, Fernando
APPLICANT: KROLL, Mathias
APPLICANT: CONDORET, Jean-Paul
TITLE OF INVENTION: Human beta-TrCP protein
FILE REFERENCE: 935.38812X00
CURRENT APPLICATION NUMBER: US/09/601,168B
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: PCT/FR99/00196
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: FR98 01100
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: FR98 15545
PRIOR FILING DATE: 1998-12-09
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1 and manually
SEQ ID NO 2
LENGTH: 569
TYPE: prt
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Description of the artificial sequence : CDNA
OTHER INFORMATION: coding for human beta-TrCP protein
US-09-601-168B-2

Query Match 100.0% Score 3034; DB 5; Length 569;
Best Local Similarity 100.0%; Pred No. 5e-273;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPRAVLOEKALKFNMSSEREDCNNGEPKRIIPKNSLRQYNSCARLCAQETVCA 60
1 MDPRAVLOEKALKFNMSSEREDCNNGEPKRIIPKNSLRQYNSCARLCAQETVCA 60
61 STAKTENCYAKTKLANGTSMTVPKOKLSASYEKEKELCVYFFQWSSDVEFEVHL 120
61 STAKTENCYAKTKLANGTSMTVPKOKLSASYEKEKELCVYFFQWSSDVEFEVHL 120
121 ISQCHQOHGHINSYLPMLORDEFTALPARGDHIAENILSYLDKSLCAELVCKEY 180
121 ISQCHQOHGHINSYLPMLORDEFTALPARGDHIAENILSYLDKSLCAELVCKEY 180
121 ISQCHQOHGHINSYLPMLORDEFTALPARGDHIAENILSYLDKSLCAELVCKEY 180
181 RVTSQMLMKLIERVVRTSLMGAEERRGQGYLFKNPDPGNAAPNSFYALPKII 240
181 RVTSQMLMKLIERVVRTSLMGAEERRGQGYLFKNPDPGNAAPNSFYALPKII 240

```

Db 181 RYSDGMLMKLIERVNTSDLMKGLAERRGQYLFFKNKPPDGNAPNSFYALYKII 240
QY 241 QDIETISNMRCGRHSIORIHCHSETSKGYCCLOYDDOKIVSGLRNTIKIMDKNTLECK 300
Db 241 QDIETISNMRCGRHSIORIHCHSETSKGYCCLOYDDOKIVSGLRNTIKIMDKNTLECK 300
QY 301 RILGHTGVSVCLOYDERVITITSSSDTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
Db 301 RILGHTGVSVCLOYDERVITITSSSDTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
QY 361 VTCSKDSIAVMDASPTDITLRVILVGHRAAVNVDPDDKTVASGORTIKVWNTSIC 420
Db 361 VTCSKDSIAVMDASPTDITLRVILVGHRAAVNVDPDDKTVASGORTIKVWNTSIC 420
QY 421 EFVFTLNGHRGIAICLOYDRILVSSGSDNTYRLMDIEGACILVLEGEHELVRCIRPDN 480
Db 421 EFVFTLNGHRGIAICLOYDRILVSSGSDNTYRLMDIEGACILVLEGEHELVRCIRPDN 480
QY 481 KRIYSGAYDGKIKVMDLVALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540
Db 481 KRIYSGAYDGKIKVMDLVALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540
QY 541 ILIMDFLNDPAQAEPSPRSRTTYISR 569
Db 541 ILIMDFLNDPAQAEPSPRSRTTYISR 569

```

RESULT 2

```

US-09-724-676-52731
; Sequence 52731, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52731
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-52731

```

Query Match 99.1%; Score 3006; DB 5; Length 605;
Best Local Similarity 94.0%; Pred. No. 2.2e-270;

Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

```

QY 1 MDPAEAVILOEKALKFM-----NSSEREDC 24
Db 1 MDPAEAVILOEKALKFMCSMPRLMLGSSSLADSMPSILRCIYNPGGALITAFQNSSEREDC 60
QY 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAKTKLANGTSSMIV 84
Db 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAKTKLANGTSSMIV 120
QY 85 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 144
Db 85 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 180
QY 121 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 180
Db 121 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 180
QY 145 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 204
Db 145 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 204
QY 181 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 240
Db 181 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 240
QY 205 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 264
Db 205 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 300
QY 241 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 300
Db 241 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 300
QY 265 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 324
Db 265 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 360
QY 301 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 360
Db 301 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 360
QY 325 SDSTYRWMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTSKDSIAVMDASPTDITLRR 384

```

```

Db 361 SDSTYRWMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTSKDSIAVMDASPTDITLRR 420
QY 385 VLVGHRAAVNVVDEDDKTVIASAGDRTIKVWNTSTCEPFTLNGHRGIAICLOYDRILV 444
Db 421 VLVGHRAAVNVVDEDDKTVIASAGDRTIKVWNTSTCEPFTLNGHRGIAICLOYDRILV 480
QY 445 SGSSDNTIRLMDIEGACILVLEGEHELVRCIRPDNKRIVSGAYDGKIKVMDLVALDPR 504
Db 481 SGSSDNTIRLMDIEGACILVLEGEHELVRCIRPDNKRIVSGAYDGKIKVMDLVALDPR 540
QY 505 APAGTLCRTLVHSGRVRLQDFEQIVSSSHDTILIMDFLNDPAQAEPSPRSRTY 564
Db 541 APAGTLCRTLVHSGRVRLQDFEQIVSSSHDTILIMDFLNDPAQAEPSPRSRTY 600
QY 565 TYISR 569
Db 601 TYISR 605

```

RESULT 3

```

US-09-724-676A-52731
; Sequence 52731, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52731
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-52731

```

Query Match 99.1%; Score 3006; DB 5; Length 605;
Best Local Similarity 94.0%; Pred. No. 2.2e-270;

Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

```

QY 1 MDPAEAVILOEKALKFM-----NSSEREDC 24
Db 1 MDPAEAVILOEKALKFMCSMPRLMLGSSSLADSMPSILRCIYNPGGALITAFQNSSEREDC 60
QY 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAKTKLANGTSSMIV 84
Db 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAKTKLANGTSSMIV 120
QY 85 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 144
Db 85 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 180
QY 121 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 180
Db 121 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISOMCHYOHGHINSYLPMLQDRF 180
QY 145 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 204
Db 145 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 240
QY 181 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 240
Db 181 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMLMKLIERVNTDSLNR 240
QY 205 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 264
Db 205 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 300
QY 241 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 300
Db 241 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETISNMRCGRHSIORIHCHS 300
QY 265 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 324
Db 265 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 360
QY 301 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 360
Db 301 ETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITGS 360
QY 325 SDSTYRWMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTSKDSIAVMDASPTDITLRR 384
Db 361 SDSTYRWMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTSKDSIAVMDASPTDITLRR 420
QY 385 VLVGHRAAVNVVDEDDKTVIASAGDRTIKVWNTSTCEPFTLNGHRGIAICLOYDRILV 444
Db 421 VLVGHRAAVNVVDEDDKTVIASAGDRTIKVWNTSTCEPFTLNGHRGIAICLOYDRILV 480

```

OY 445 SSSDNTIRLMDIEGACIRVLEGHEELVRCIRPDNKRIVSGADGKIKVMDVLAALDPR 504
 |||||
 Db 481 SSSDNTIRLMDIEGACIRVLEGHEELVRCIRPDNKRIVSGADGKIKVMDVLAALDPR 540
 OY 505 APAGTICLRTVEHSGVRVLOFDEFQIVSSSHDITILIMFLNDPAAQAEPPSPSKTY 564
 |||||
 Db 541 APAGTICLRTVEHSGVRVLOFDEFQIVSSSHDITILIMFLNDPAAQAEPPSPSKTY 600
 OY 565 TYISR 569
 |||||
 Db 601 TYISR 605

RESULT 4
 US-09-724-676-52729
 ; Sequence 52729, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; NUMBER OF FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 52729
 ; LENGTH: 537
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-52729

Query Match 89.0%; Score 2701; DB 5; Length 537;
 Best Local Similarity 91.2%; Pred. No. 3.8e-242;

Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

OY 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLNOETVCLA 60
 |||||
 Db 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLNOETVCLA 60
 OY 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKEKCYKFEQWESDQVEFEHL 120
 |||||
 Db 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKEKCYKFEQWESDQVEFEHL 120
 OY 121 ISOMCHYOHGINSYKPKMLQRODFTALPARGLDHAENILSYLDAKSLCAAEVCKEY 180
 |||||
 Db 121 ISOMCHYOHGINSYKPKMLQRODFTALPARGLDHAENILSYLDAKSLCAAEVCKEY 180
 OY 181 RYSDGMIMKLLERAVRTDSLMGRLAERKMGQYLFKNRPPDGNAPNSFYALYPKII 240
 |||||
 Db 181 RYSDGMIMKLLERAVRTDSLMGRLAERKMGQYLFKNRPPDGNAPNSFYALYPKII 240
 OY 241 ODIEETIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300
 |||||
 Db 241 ODIEETIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300
 OY 213 ---TIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268
 |||||
 Db 213 ---TIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268
 OY 301 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 |||||
 Db 301 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 OY 269 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 328
 |||||
 Db 269 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 328
 OY 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDFDDKYIYASAGDRTIKVWNTSTC 420
 |||||
 Db 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDFDDKYIYASAGDRTIKVWNTSTC 420
 OY 421 EFVRLNGHKRGIAQLQYRDRLVYVSGSSDNTIRLMDIEGACIRVLEGHEELVRCIRPDN 480
 |||||
 Db 421 EFVRLNGHKRGIAQLQYRDRLVYVSGSSDNTIRLMDIEGACIRVLEGHEELVRCIRPDN 480
 OY 481 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 540
 |||||
 Db 481 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 540
 OY 449 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 508
 |||||
 Db 449 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 508
 OY 541 ILIMFLNDPAAQAEPPSPSKTYISR 569
 |||||
 Db 541 ILIMFLNDPAAQAEPPSPSKTYISR 569
 OY 509 ILIMFLNDPAAQAEPPSPSKTYISR 537
 |||||

RESULT 5
 US-09-724-676-52730
 ; Sequence 52730, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; NUMBER OF FILING DATE: 2000-11-28
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 52730
 ; LENGTH: 537
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-52730

Query Match 89.0%; Score 2701; DB 5; Length 537;
 Best Local Similarity 91.2%; Pred. No. 3.8e-242;

Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

OY 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLNOETVCLA 60
 |||||
 Db 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLNOETVCLA 60
 OY 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKEKCYKFEQWESDQVEFEHL 120
 |||||
 Db 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKEKCYKFEQWESDQVEFEHL 120
 OY 121 ISOMCHYOHGINSYKPKMLQRODFTALPARGLDHAENILSYLDAKSLCAAEVCKEY 180
 |||||
 Db 121 ISOMCHYOHGINSYKPKMLQRODFTALPARGLDHAENILSYLDAKSLCAAEVCKEY 180
 OY 181 RYSDGMIMKLLERAVRTDSLMGRLAERKMGQYLFKNRPPDGNAPNSFYALYPKII 240
 |||||
 Db 181 RYSDGMIMKLLERAVRTDSLMGRLAERKMGQYLFKNRPPDGNAPNSFYALYPKII 240
 OY 241 ODIEETIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300
 |||||
 Db 241 ODIEETIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300
 OY 213 ---TIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268
 |||||
 Db 213 ---TIESNMGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268
 OY 301 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 |||||
 Db 301 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 OY 269 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 328
 |||||
 Db 269 RILTHTGTSVCLQYDEVYITGSSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGM 328
 OY 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDFDDKYIYASAGDRTIKVWNTSTC 420
 |||||
 Db 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDFDDKYIYASAGDRTIKVWNTSTC 420
 OY 421 EFVRLNGHKRGIAQLQYRDRLVYVSGSSDNTIRLMDIEGACIRVLEGHEELVRCIRPDN 480
 |||||
 Db 421 EFVRLNGHKRGIAQLQYRDRLVYVSGSSDNTIRLMDIEGACIRVLEGHEELVRCIRPDN 480
 OY 481 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 540
 |||||
 Db 481 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 540
 OY 449 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 508
 |||||
 Db 449 KRIVSAYGKIKVMDVLAALDPRAPAGTICLRTVEHSGVRVLOFDEFQIVSSSHDT 508
 OY 541 ILIMFLNDPAAQAEPPSPSKTYISR 569
 |||||
 Db 541 ILIMFLNDPAAQAEPPSPSKTYISR 569
 OY 509 ILIMFLNDPAAQAEPPSPSKTYISR 537
 |||||

RESULT 6
 US-09-724-676A-52729
 ; Sequence 52729, Application US/09724676A

; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222
 SOFTWARE: Patentin version 3.2
 SEQ ID NO 52729
 LENGTH: 537
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-724-676A-52732

Query Match 89.0%; Score 2701; DB 5; Length 537;
 Best Local Similarity 91.2%; Pred. No. 3.8e-242;
 Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

```

QY 1 MDPAAVLOEKALFKFNNSSEREDCNGGPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
DB 1 MDPAAVLOEKALFKFNNSSEREDCNGGPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
QY 61 STAMKTEVCYAKTKRLANGSSMIVPQKRLSASYEKEKELCYKFEQWSESQVEFEVHL 120
DB 61 STAMKTEVCYAKTKRLANGSSMIVPQKRLSASYEKEKELCYKFEQWSESQVEFEVHL 120
QY 121 ISOMCHYOHGHINSYLYKPMLOQDFITLALPARGLDHAENILSYLDKSLCAAEVCKEM 180
DB 121 ISOMCHYOHGHINSYLYKPMLOQDFITLALPARGLDHAENILSYLDKSLCAAEVCKEM 180
QY 181 RVTSQDMLMKKLIERNVTRDLSMRGLAERGGQYLFKNKPPDGNAPNSRYALYPKII 240
DB 181 RVTSQDMLMKKLIERNVTRDLSMRGLAERGGQYLFKNKPPDGNAPNSRYALYPKII 240
QY 241 QDIETTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
DB 241 QDIETTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
QY 213 ---TTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 268
DB 213 ---TTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 268
QY 301 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 360
DB 301 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 360
QY 269 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 328
DB 269 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 328
QY 361 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 420
DB 361 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 420
QY 329 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 388
DB 329 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 388
QY 421 EFRVRLNGHKKRIACLOYRDRLVYSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
DB 421 EFRVRLNGHKKRIACLOYRDRLVYSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
QY 481 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 540
DB 481 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 540
QY 541 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 508
DB 541 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 508
QY 509 ILIMDFLMDPAAQAEPSPSRITTYISR 537
DB 509 ILIMDFLMDPAAQAEPSPSRITTYISR 537

```

RESULT 7

US-09-724-676A-52730
 Sequence 52730, Application US/09724676A
 GENERAL INFORMATION:
 APPLICANT: Compugen LTD
 TITLE OF INVENTION: Variants of alternative splicing
 FILE REFERENCE: 129181.4 Compugen
 CURRENT APPLICATION NUMBER: US/09/724,676A
 NUMBER OF SEQ ID NOS: 97222
 SOFTWARE: Patentin version 3.2
 SEQ ID NO 52730
 LENGTH: 537
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-724-676A-52730

Query Match 89.0%; Score 2701; DB 5; Length 537;
 Best Local Similarity 91.2%; Pred. No. 3.8e-242;
 Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

```

QY 1 MDPAAVLOEKALFKFNNSSEREDCNGGPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
DB 1 MDPAAVLOEKALFKFNNSSEREDCNGGPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
QY 61 STAMKTEVCYAKTKRLANGSSMIVPQKRLSASYEKEKELCYKFEQWSESQVEFEVHL 120
DB 61 STAMKTEVCYAKTKRLANGSSMIVPQKRLSASYEKEKELCYKFEQWSESQVEFEVHL 120
QY 121 ISOMCHYOHGHINSYLYKPMLOQDFITLALPARGLDHAENILSYLDKSLCAAEVCKEM 180
DB 121 ISOMCHYOHGHINSYLYKPMLOQDFITLALPARGLDHAENILSYLDKSLCAAEVCKEM 180
QY 181 RVTSQDMLMKKLIERNVTRDLSMRGLAERGGQYLFKNKPPDGNAPNSRYALYPKII 240
DB 181 RVTSQDMLMKKLIERNVTRDLSMRGLAERGGQYLFKNKPPDGNAPNSRYALYPKII 240
QY 241 QDIETTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
DB 241 QDIETTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
QY 213 ---TTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 268
DB 213 ---TTESNMRCGRSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECK 268
QY 301 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 360
DB 301 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 360
QY 269 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 328
DB 269 RLTGHTGSLVLCQYDERVITIGSSDSTVRYWDVNTGEMNTLLIHCEAVLHLRFNNGM 328
QY 361 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 420
DB 361 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 420
QY 329 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 388
DB 329 VTCSDKRSIAVMDMSPDITLRLVYGHRAAVNVVDEDDKYIVASGDRITIKWNTSTC 388
QY 421 EFRVRLNGHKKRIACLOYRDRLVYSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
DB 421 EFRVRLNGHKKRIACLOYRDRLVYSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
QY 481 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 540
DB 481 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 540
QY 541 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 508
DB 541 KRIYGAVDGKIKYWDVLAALDPAPAGTLCRLTVHSGVRFLOQDFEQIVSSSHDT 508
QY 509 ILIMDFLMDPAAQAEPSPSRITTYISR 537
DB 509 ILIMDFLMDPAAQAEPSPSRITTYISR 537

```

RESULT 8

US-09-724-676-52732
 Sequence 52732, Application US/09724676
 GENERAL INFORMATION:
 APPLICANT: Compugen LTD
 TITLE OF INVENTION: Variants of alternative splicing
 FILE REFERENCE: 129181.4 Compugen
 CURRENT APPLICATION NUMBER: US/09/724,676
 NUMBER OF SEQ ID NOS: 97222
 SOFTWARE: Patentin version 3.2
 SEQ ID NO 52732
 LENGTH: 573
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-724-676-52732

Query Match 88.1%; Score 2673; DB 5; Length 573;
 Best Local Similarity 85.8%; Pred. No. 1.7e-239;
 Matches 519; Conservative 4; Mismatches 14; Indels 68; Gaps 3;

```

QY 1 MDPAAVLOEKALFKFNNSSEREDCNGGPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
DB 1 MDPAAVLOEKALFKFNNSSEREDCNGGPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
QY 25 NNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLASTAMKTEVCYAKTKLANGTSSMIV 84
DB 61 NNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLASTAMKTEVCYAKTKLANGTSSMIV 120
QY 85 PKORKLSASYEKEKELCYKFEQWSESQVEFEVHLISOMCHYOHGHINSYLYKPMLOQDF 144
DB 85 PKORKLSASYEKEKELCYKFEQWSESQVEFEVHLISOMCHYOHGHINSYLYKPMLOQDF 144

```



```

Db 121 PKORKLSASYEKELCVYFEQWSESDOYEVEHLISQMHYQGHINSYLMQORDF 180
QY 145 ITALPARGLDHAENILSYLDAKSLCAELVCKEYRVTSQDMKMLIERVVRTDSLNR 204
Db 181 ITALPARGLDHAENILSYLDAKSLCAELVCKEYRVTSQDMKMLIERVVRTDSLNR 236
QY 205 GLAERGMGOYLFRKKNRPPDGNAPNSFYRALYPKIIQDIETIESNMRCGRHSIORIHCRS 264
Db 237 FSEVERGRTKRM-----TIESNMRCGRHSIORIHCRS 268
QY 265 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSVLCLOYDERVITIGS 324
Db 269 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSVLCLOYDERVITIGS 328
QY 325 SDSTVAVMDVNTGEMNLTIHCEAVLHLRFNNGMMVTCSDRSIAVWDMASPTDITLRR 384
Db 329 SDSTVAVMDVNTGEMNLTIHCEAVLHLRFNNGMMVTCSDRSIAVWDMASPTDITLRR 388
QY 385 VLVGHRAAVNVVDFDRTKYVSASGDRITKVMNTSTCEVFRTLNGHKGRIACLOYRDLVY 444
Db 389 VLVGHRAAVNVVDFDRTKYVSASGDRITKVMNTSTCEVFRTLNGHKGRIACLOYRDLVY 448
QY 445 SSSSDNTIRLMDIEGACLRVLEGEHELVRCLRFDNKRIVSGAIDGKIKVMDLVAAALDPR 504
Db 449 SSSSDNTIRLMDIEGACLRVLEGEHELVRCLRFDNKRIVSGAIDGKIKVMDLVAAALDPR 508
QY 505 APAGTLCRLTIVESHSGRVFLQDFEFQIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 564
Db 509 APAGTLCRLTIVESHSGRVFLQDFEFQIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 568
QY 565 TYISR 569
Db 569 TYISR 573

RESULT 9
US-09-724-676-52733
; Sequence 52733, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52733
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-52733

Query Match      88.1%; Score 2673; DB 5; Length 573;
Best Local Similarity 85.8%; Pred. No. 1.7e-239;
Matches 519; Conservative 4; Mismatches 14; Indels 68; Gaps 3;

```

```

Db 237 FSEVERGRTKRM-----TIESNMRCGRHSIORIHCRS 268
QY 265 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSVLCLOYDERVITIGS 324
Db 269 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSVLCLOYDERVITIGS 328
QY 325 SDSTVAVMDVNTGEMNLTIHCEAVLHLRFNNGMMVTCSDRSIAVWDMASPTDITLRR 384
Db 329 SDSTVAVMDVNTGEMNLTIHCEAVLHLRFNNGMMVTCSDRSIAVWDMASPTDITLRR 388
QY 385 VLVGHRAAVNVVDFDRTKYVSASGDRITKVMNTSTCEVFRTLNGHKGRIACLOYRDLVY 444
Db 389 VLVGHRAAVNVVDFDRTKYVSASGDRITKVMNTSTCEVFRTLNGHKGRIACLOYRDLVY 448
QY 445 SSSSDNTIRLMDIEGACLRVLEGEHELVRCLRFDNKRIVSGAIDGKIKVMDLVAAALDPR 504
Db 449 SSSSDNTIRLMDIEGACLRVLEGEHELVRCLRFDNKRIVSGAIDGKIKVMDLVAAALDPR 508
QY 505 APAGTLCRLTIVESHSGRVFLQDFEFQIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 564
Db 509 APAGTLCRLTIVESHSGRVFLQDFEFQIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 568
QY 565 TYISR 569
Db 569 TYISR 573

RESULT 10
US-09-724-676A-52732
; Sequence 52732, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52732
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-52732

Query Match      88.1%; Score 2673; DB 5; Length 573;
Best Local Similarity 85.8%; Pred. No. 1.7e-239;
Matches 519; Conservative 4; Mismatches 14; Indels 68; Gaps 3;

```

QY	325	SDSIYRVMDVNTGEMLNTLIHHCEAVLHLFRNNCMNTYCSKDRSTIAWMDASPPDITLRR	364
Db	329	SDSIYRVMDVNTGEMLNTLIHHCEAVLHLFRNNCMNTYCSKDRSTIAWMDASPPDITLRR	368
QY	365	VLVGHRAAVNVVDEDDKIYVSASDRTIKWNTSTCFEVRTLNGHKRGIACTOYRDLVY	444
Db	369	VLVGHRAAVNVVDEDDKIYVSASDRTIKWNTSTCFEVRTLNGHKRGIACTOYRDLVY	448
QY	445	SGSSDNTIRLMDIECGACLRVLEHHEELVRCIRFDNKRIVSGADGKIKWDLVLAALDPR	504
Db	449	SGSSDNTIRLMDIECGACLRVLEHHEELVRCIRFDNKRIVSGADGKIKWDLVLAALDPR	508
QY	505	APAGTLCRLTVEHSGRVRLQPDDEFQIVSSSHDDTLILMDFLNDPAPAQAEPSPSPRTY	564
Db	509	APAGTLCRLTVEHSGRVRLQPDDEFQIVSSSHDDTLILMDFLNDPAPAQAEPSPSPRTY	568
QY	565	TYISR 569	
Db	569	TYISR 573	

RESULT 11
US-09-724-676A-52733
; Sequence 52733, Application US/09724676A

Query Match	88.1%	Score	2673	DB	5	Length	573
Best Local Similarity	85.8%	Pred	No. 1.7e-239				
Matches	519	Conservative	4	Mismatches	14	Indels	68
						Gaps	3

```

QY      445  SSSSSNTIRLMDIEGACLARLYEGHEELVRCIRFDNKRIVSGAADGAKITWDLVAAALDPR 504
Db      449  SSSSSNTIRLMDIEGACLARLYEGHEELVRCIRFDNKRIVSGAADGAKITWDLVAAALDPR 508
QY      505  APAGTLCRLTVEHSGRVRFLQFDEFOIVSSSHDFTILIMDFLNDPAAQAEPPRPSRTY 564
Db      509  APAGTLCRLTVEHSGRVRFLQFDEFOIVSSSHDFTILIMDFLNDPAAQAEPPRPSRTY 568
QY      565  TYISR 569
Db      569  TYISR 573

RESULT 12
US-09-724-676-61015
; Sequence 61015, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61015
; LENGTH: 529
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-61015

```

Query Match	79.1%;	Score 2400;	DB 5;	Length 529;
Best Local Similarity	79.3%;	Pred. No. 3.3e-214;		
Matches 451;	Conservative 43;	Mismatches 35;	Indels 40;	Gaps 4

Matches	451;	Conservative	43;	Mismatches	35;	Indels	40;	Gaps	4;
---------	------	--------------	-----	------------	-----	--------	-----	------	----

QY	1	MDPAEVLIOEKLKATKMNSENBERDCNNGBPBKIIIPKNSLJQOTYNSCARLCLNOEYCLA	60
Db	1	MEP-DSVIEDKTLIELMNTSVEMEDONEDSPK-----KNTLMQ-----	36
QY	61	STAMKTENCVAATKLANOTSSMIYKPKOKLSASYEKEKEKELCVKFFEQMSSEDQVEFVHL	120
Db	37	-----ISNCTSSYIVSRKRPBEGCNOKCKEDLCIKFFDOMSESDQVEFVHL	82
QY	121	ISOMCHYOHGHINSLYKFLPMLORDFTALPANGLDHIAENILSYLDANSILCAELVCKEMY	180
Db	83	ISRMCHYOHGHINSLYKFLPMLORDFTALPEOGLDHIAENILSYLDANSILCAELVCKEMQ	144
QY	181	RYTSGMLMKKLTLEMVYTDLSMRLAARBCMGOLYFRNKPDDGNAPPNSRYALYKII	240
Db	143	RYISGGMMLMKKLTLEMVYTDLPMLGLSRRKMDOLYFNKRPDQ--PNSFYRSLYKPII	200
QY	241	QDIETIESNMRCGRHSIORICRSETSKGYVCLQYDDOKVYSGLRDNTIKTMDNTECK	300
Db	201	QDIETIESNMRCGRHNIQICRSRNSKGYVCLQYDDEKIIISGLRDSIKTMDTSLIECL	260
QY	301	RILGHTGTSVLCLOYDERVYIITGSSDSTVRWVDVNTGEMLNTLILHCEAVLHLRFNNGM	360
Db	261	KYLGHITSVLCLOYDERVYITGSSDSTVRWVDVNTGELVTLTLIHNEAVLHLRFNSGLM	320
QY	361	WYCSKDRSIAWDMASPDITLIRLRYLVCHRAAVVWVPDDKYIVSASGDRIKYKWNSTC	420
Db	321	WYCSKDRSIAWDMASADITLIRLRYLVCHRAAVVWVPDDKYIVSASGDRIKYKWNSTC	380
QY	421	EFVRTLNGHKRGICLQYRDLRYVSSGSDNTRIIMDJECGACILVLEGHEELVYCIRFDN	480
Db	381	EFVRTLNGHKRGICLQYRDLRYVSSGSDNTRIIMDJECGACILVLEGHEELVYRCIFPDN	440
QY	481	KRIYSGAYDGKIKYWDLYAALDPAAPAGTCLRTLVHSGHVRFLQDEFOIYSSHDDT	540
Db	441	KRIYSGAYDGKIKYWDLOAALDPAAPASTCLRTLVHSGHVRFLQDEFOIYSSHDDT	500
QY	541	ILIMDFLNDPAQAQAEPPRSPRRTYISR	569
Db	501	ILIMDFLNDPAQAQAEPPRSPRRTYISR	529

```
Search completed: February 20, 2003, 09:59:48
Job time : 59 secs
```


Db 61 STAKTEKCAKTLKANGTSSMIVPKOKLSASTEKEKELCVKFFEDWSSDVEFEHL 120
QY 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDKSLCAELVCKEMY 180
Db 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDKSLCAELVCKEMY 180
QY 181 RYSDGMLMKKLLIERMVRTDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240
Db 181 RYSDGMLMKKLLIERMVRTDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240
QY 241 ODITISNMRCGRHSLOIRHCRSETSKGYVCLQYDDOKIVSGLDNNTIKIMDKNTLECK 300
Db 241 ODITISNMRCGRHSLOIRHCRSETSKGYVCLQYDDOKIVSGLDNNTIKIMDKNTLECK 300
QY 301 RILGHTGSVLCLOYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRPNNGM 360
Db 301 RILGHTGSVLCLOYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRPNNGM 360
QY 361 VTCSKDRSIAVWDMASPTDITLRLVLYGHRAAVNVVDEDDKYIVSASGDRITKWNSTC 420
Db 361 VTCSKDRSIAVWDMASPTDITLRLVLYGHRAAVNVVDEDDKYIVSASGDRITKWNSTC 420
QY 421 EFWRLNGHKGKIGIACLOYRDLVYSGSSDNTIRLMDIEGACLRVLEHHELVRCIRFDN 480
Db 421 EFWRLNGHKGKIGIACLOYRDLVYSGSSDNTIRLMDIEGACLRVLEHHELVRCIRFDN 480
QY 481 KRIVSAGYDGKIKWMDIYVALDPRAPAGTICLRTLVHSGRVFRLQDFEFOIVSSHDDT 540
Db 481 KRIVSAGYDGKIKWMDIYVALDPRAPAGTICLRTLVHSGRVFRLQDFEFOIVSSHDDT 540
QY 541 ILIWFNDLNDPAQAOPRPSRRTTYISR 569
Db 541 ILIWFNDLNDPAQAOPRPSRRTTYISR 569

RESULT 2
PCT-US99-19560-2
Sequence 2, Application PC/TUS9919560
GENERAL INFORMATION:
APPLICANT: NEW YORK UNIVERSITY
TITLE OF INVENTION: NOVEL DBIQUITTIN LIGASES AS THERAPEUTIC TARGETS
FILE REFERENCE: 5914-081-228
CURRENT APPLICATION NUMBER: PCT/US99/19560
EARLIER FILING DATE: 1999-08-31
EARLIER APPLICATION NUMBER: 60/098,355
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: 60/118,568
EARLIER FILING DATE: 1999-02-03
EARLIER APPLICATION NUMBER: 60/124,449
EARLIER FILING DATE: 1999-03-15
NUMBER OF SEQ ID NOS: 89
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 569
TYPE: PRF
ORGANISM: Homo sapiens
PCT-US99-19560-2

Query Match 100.0%; Score 3034; DB 1; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDKSLCAELVCKEMY 180
QY 181 RYSDGMLMKKLLIERMVRTDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240
Db 181 RYSDGMLMKKLLIERMVRTDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240
QY 241 ODITISNMRCGRHSLOIRHCRSETSKGYVCLQYDDOKIVSGLDNNTIKIMDKNTLECK 300
Db 241 ODITISNMRCGRHSLOIRHCRSETSKGYVCLQYDDOKIVSGLDNNTIKIMDKNTLECK 300
QY 301 RILGHTGSVLCLOYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRPNNGM 360
Db 301 RILGHTGSVLCLOYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRPNNGM 360
QY 361 VTCSKDRSIAVWDMASPTDITLRLVLYGHRAAVNVVDEDDKYIVSASGDRITKWNSTC 420
Db 361 VTCSKDRSIAVWDMASPTDITLRLVLYGHRAAVNVVDEDDKYIVSASGDRITKWNSTC 420
QY 421 EFWRLNGHKGKIGIACLOYRDLVYSGSSDNTIRLMDIEGACLRVLEHHELVRCIRFDN 480
Db 421 EFWRLNGHKGKIGIACLOYRDLVYSGSSDNTIRLMDIEGACLRVLEHHELVRCIRFDN 480
QY 481 KRIVSAGYDGKIKWMDIYVALDPRAPAGTICLRTLVHSGRVFRLQDFEFOIVSSHDDT 540
Db 481 KRIVSAGYDGKIKWMDIYVALDPRAPAGTICLRTLVHSGRVFRLQDFEFOIVSSHDDT 540
QY 541 ILIWFNDLNDPAQAOPRPSRRTTYISR 569
Db 541 ILIWFNDLNDPAQAOPRPSRRTTYISR 569

RESULT 3
US-09-210-060-18
Sequence 18, Application US/09210060
GENERAL INFORMATION:
APPLICANT: Manning, Anthony M.
APPLICANT: Mercurio, Frank
APPLICANT: Amlt, Sharon
APPLICANT: Ben-Neriah, Yimon
APPLICANT: Davis, Matti
APPLICANT: Hatzubai, Ada
APPLICANT: Lavyon, Itis
APPLICANT: Yaron, Avraham
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
TITLE OF INVENTION: NF-kB
FILE REFERENCE: 860098.427
CURRENT APPLICATION NUMBER: US/09/210.060
EARLIER FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 18
LENGTH: 569
TYPE: PRF
ORGANISM: Homo sapiens
US-09-210-060-18

Query Match 100.0%; Score 3034; DB 16; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLOERALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLINOETVCLA 60
Db 1 MDPAAVLOERALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLINOETVCLA 60
QY 61 STAKTEKCAKTLKANGTSSMIVPKOKLSASTEKEKELCVKFFEDWSSDVEFEHL 120
Db 61 STAKTEKCAKTLKANGTSSMIVPKOKLSASTEKEKELCVKFFEDWSSDVEFEHL 120
QY 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDKSLCAELVCKEMY 180
Db 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDKSLCAELVCKEMY 180

QY 1 MDPAAVLOERALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLINOETVCLA 60
Db 1 MDPAAVLOERALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARCLINOETVCLA 60
QY 61 STAKTEKCAKTLKANGTSSMIVPKOKLSASTEKEKELCVKFFEDWSSDVEFEHL 120
Db 61 STAKTEKCAKTLKANGTSSMIVPKOKLSASTEKEKELCVKFFEDWSSDVEFEHL 120
QY 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDKSLCAELVCKEMY 180
Db 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDKSLCAELVCKEMY 180
QY 181 RYSDGMLMKKLLIERMVRTDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240
Db 181 RYSDGMLMKKLLIERMVRTDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240

```

Db 181 RVTSDGMLMKKLLERVRRTDLSLRGLAERRGQYLFKKKPPDGNAPNSFYRALYPKII 240
QY 241 ODIEETESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
Db 241 ODIEETESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
QY 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMNLTLHHCBAVHLRFNNGM 360
Db 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMNLTLHHCBAVHLRFNNGM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIYVSASGDRITIKWNTSTC 420
Db 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIYVSASGDRITIKWNTSTC 420
QY 421 EFRVLNCHKRGKGIACQYDRDLVYVSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
Db 421 EFRVLNCHKRGKGIACQYDRDLVYVSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVHSGRVFRLQDFEQIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVHSGRVFRLQDFEQIVSSSHDT 540
QY 541 ILIWDPLNDPAAQAEPSPRSRTTYISR 569
Db 541 ILIWDPLNDPAAQAEPSPRSRTTYISR 569

```

RESULT 4

```

US-09-385-219-2
; Sequence 2, Application US/09385219
; GENERAL INFORMATION:
; APPLICANT: Chlauro, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/09/385,219
; EARLIER FILING DATE: 1999-08-27
; EARLIER APPLICATION NUMBER: 60/098,355
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/118,568
; EARLIER FILING DATE: 1999-02-03
; EARLIER APPLICATION NUMBER: 60/124,449
; EARLIER FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-385-219-2

```

Query Match 100.0%; Score 3034; DB 17; Length 569;

Best Local Similarity 100.0%; Pred. No. 1.3e-284; Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDPAAVLOEKALKFNNSSEREDCNGEPKRIPEKNSLRQTYNSCARCLINQETVCLA 60
Db 1 MDPAAVLOEKALKFNNSSEREDCNGEPKRIPEKNSLRQTYNSCARCLINQETVCLA 60
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLASYEKEKELCVYFEQWSESDDVEFEHL 120
Db 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLASYEKEKELCVYFEQWSESDDVEFEHL 120
QY 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDIAENITSLYDAKSLCAAEVCKEMY 180
Db 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDIAENITSLYDAKSLCAAEVCKEMY 180
QY 181 RVTSDGMLMKKLLERVRRTDLSLRGLAERRGQYLFKKKPPDGNAPNSFYRALYPKII 240
Db 181 RVTSDGMLMKKLLERVRRTDLSLRGLAERRGQYLFKKKPPDGNAPNSFYRALYPKII 240

```

```

QY 241 ODIEETESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
Db 241 ODIEETESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
QY 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMNLTLHHCBAVHLRFNNGM 360
Db 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMNLTLHHCBAVHLRFNNGM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIYVSASGDRITIKWNTSTC 420
Db 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIYVSASGDRITIKWNTSTC 420
QY 421 EFRVLNCHKRGKGIACQYDRDLVYVSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
Db 421 EFRVLNCHKRGKGIACQYDRDLVYVSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVHSGRVFRLQDFEQIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVHSGRVFRLQDFEQIVSSSHDT 540
QY 541 ILIWDPLNDPAAQAEPSPRSRTTYISR 569
Db 541 ILIWDPLNDPAAQAEPSPRSRTTYISR 569

```

RESULT 5

```

US-09-415-795-4
; Sequence 4, Application US/09415795
; GENERAL INFORMATION:
; APPLICANT: Zhou, Pengbo
; APPLICANT: Howley, Peter M.
; TITLE OF INVENTION: TARGETED PROTEOLYSIS BY RECRUITMENT TO UBIQUITIN
; TITLE OF INVENTION: PROTEIN LIGASES
; FILE REFERENCE: HMV-043.01
; CURRENT APPLICATION NUMBER: US/09/415,795
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-415-795-4

```

Query Match 100.0%; Score 3034; DB 18; Length 569;

Best Local Similarity 100.0%; Pred. No. 1.3e-284; Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDPAAVLOEKALKFNNSSEREDCNGEPKRIPEKNSLRQTYNSCARCLINQETVCLA 60
Db 1 MDPAAVLOEKALKFNNSSEREDCNGEPKRIPEKNSLRQTYNSCARCLINQETVCLA 60
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLASYEKEKELCVYFEQWSESDDVEFEHL 120
Db 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLASYEKEKELCVYFEQWSESDDVEFEHL 120
QY 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDIAENITSLYDAKSLCAAEVCKEMY 180
Db 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDIAENITSLYDAKSLCAAEVCKEMY 180
QY 181 RVTSDGMLMKKLLERVRRTDLSLRGLAERRGQYLFKKKPPDGNAPNSFYRALYPKII 240
Db 181 RVTSDGMLMKKLLERVRRTDLSLRGLAERRGQYLFKKKPPDGNAPNSFYRALYPKII 240
QY 241 ODIEETESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
Db 241 ODIEETESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
QY 421 EFRVLNCHKRGKGIACQYDRDLVYVSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
Db 421 EFRVLNCHKRGKGIACQYDRDLVYVSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVHSGRVFRLQDFEQIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVHSGRVFRLQDFEQIVSSSHDT 540
QY 541 ILIWDPLNDPAAQAEPSPRSRTTYISR 569
Db 541 ILIWDPLNDPAAQAEPSPRSRTTYISR 569

```

|||||
Db 361 VTGSKDRSIAVMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
Qy 421 EFVRLNGHKKRGIACTQYDRDLVYVSSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
Db 421 EFVRLNGHKKRGIACTQYDRDLVYVSSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
Qy 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFQIVSSHDDT 540
Db 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFQIVSSHDDT 540
Qy 541 ILIMDFLNDPAAQAEPSPSRRTYIISR 569
Db 541 ILIMDFLNDPAAQAEPSPSRRTYIISR 569

RESULT 6
US-09-455-371-2
; Sequence 2, Application US/09455371
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Ellledge, Stephen
; APPLICANT: Harper, J. Wade
; APPLICANT: Rolf, Mark
; APPLICANT: Strack, Peter J.
; APPLICANT: Winston, Jeffrey T.
; TITLE OF INVENTION: Regulation of I Kappa B (Ik-B) Degradation and Methods and Reagents
; FILE REFERENCE: 120541-1010
; CURRENT APPLICATION NUMBER: US/09/455, 371
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/110982
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-455-371-2

Query Match 100.0%; Score 3034; DB 18; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
Db 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
Qy 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYEQWSESDDQVEFEHL 120
Db 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYEQWSESDDQVEFEHL 120
Qy 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
Db 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
Qy 181 RYVSDGMLMKLLIERMVRTSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRKI 240
Db 181 RYVSDGMLMKLLIERMVRTSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRKI 240
Qy 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIYVGLRDNITIKIMDKNTLECK 300
Db 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIYVGLRDNITIKIMDKNTLECK 300
Qy 301 RILTGHTGSVLCLOYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILTGHTGSVLCLOYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
Qy 361 VTGSKDRSIAVMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
Db 361 VTGSKDRSIAVMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420

Qy 421 EFVRLNGHKKRGIACTQYDRDLVYVSSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
Db 421 EFVRLNGHKKRGIACTQYDRDLVYVSSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
Qy 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFQIVSSHDDT 540
Db 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFQIVSSHDDT 540
Qy 541 ILIMDFLNDPAAQAEPSPSRRTYIISR 569
Db 541 ILIMDFLNDPAAQAEPSPSRRTYIISR 569

RESULT 7
US-09-455-371B-2
; Sequence 2, Application US/09455371B
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Ellledge, Stephen
; APPLICANT: Harper, J. Wade
; APPLICANT: Rolf, Mark
; APPLICANT: Strack, Peter J.
; APPLICANT: Winston, Jeffrey T.
; TITLE OF INVENTION: Regulation of I Kappa B (Ik-B) Degradation and Methods and Reagents
; FILE REFERENCE: 120541-1010
; CURRENT APPLICATION NUMBER: US/09/455, 371B
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/110982
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-455-371B-2

Query Match 100.0%; Score 3034; DB 18; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
Db 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
Qy 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYEQWSESDDQVEFEHL 120
Db 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYEQWSESDDQVEFEHL 120
Qy 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
Db 121 ISOMCHYOHGHINSYLLKPMLODFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
Qy 181 RYVSDGMLMKLLIERMVRTSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRKI 240
Db 181 RYVSDGMLMKLLIERMVRTSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRKI 240
Qy 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIYVGLRDNITIKIMDKNTLECK 300
Db 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIYVGLRDNITIKIMDKNTLECK 300
Qy 301 RILTGHTGSVLCLOYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILTGHTGSVLCLOYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
Qy 361 VTGSKDRSIAVMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
Db 361 VTGSKDRSIAVMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
Qy 421 EFVRLNGHKKRGIACTQYDRDLVYVSSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
Db 421 EFVRLNGHKKRGIACTQYDRDLVYVSSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480

QY 481 KRIVSAGYDKIKWDLVAALDPRAPACTLCLRTLVHSGVFRLODFEFOIVSSHDDT 540
DB 481 KRIVSAGYDKIKWDLVAALDPRAPACTLCLRTLVHSGVFRLODFEFOIVSSHDDT 540
QY 541 ILIMDFLNDPAAOAEPPRSPRSTYTYISR 569
DB 541 ILIMDFLNDPAAOAEPPRSPRSTYTYISR 569

RESULT 8

US-09-601-168a-2
; Sequence 2, Application US/09601168A
; GENERAL INFORMATION:
; APPLICANT: BENAROUS, Richard
; APPLICANT: MARGOTIN, Florence
; APPLICANT: DURAND, Hervé
; APPLICANT: ARENZANA SEISDEDOS, Fernando
; APPLICANT: KROLL, Mathias
; APPLICANT: CONDORCET, Jean-Paul
; TITLE OF INVENTION: Human beta-TTCP protein
; FILE REFERENCE: 935,38812X00
; CURRENT APPLICATION NUMBER: US/09/601,168A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: PCT/FR99/00196
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: FR98 01100
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR98 15545
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence : CDNA
; OTHER INFORMATION: coding for human beta-TTCP protein
US-09-601-168a-2

Query Match 100.0%; Score 3034; DB 20; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPBAVLOEKALFKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
DB 1 MDPBAVLOEKALFKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
QY 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFFQWSESDQVEFEVHL 120
DB 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFFQWSESDQVEFEVHL 120
QY 121 ISOMCHYOHGHINSYLPKMLORDFITLAPARGDHIENLSTYLDKSLCAALVCKEY 180
DB 121 ISOMCHYOHGHINSYLPKMLORDFITLAPARGDHIENLSTYLDKSLCAALVCKEY 180
QY 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGQYLFKNRPPGNAPNSFYALYPKII 240
DB 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGQYLFKNRPPGNAPNSFYALYPKII 240
QY 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRNDWTIKIMDKNTLECK 300
DB 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRNDWTIKIMDKNTLECK 300
QY 301 RLUTGHTGSVLCLOQDEVIITGSSDSSTVRVWDVNTGEMLNTLLHHC EAVLHLRFNNGM 360
DB 301 RLUTGHTGSVLCLOQDEVIITGSSDSSTVRVWDVNTGEMLNTLLHHC EAVLHLRFNNGM 360
QY 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420
DB 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420

QY 421 EFVRLTNGHKGRIAGLOQYRDLVWSSSDNTIRLMDIEGACALVLEGHELVRCIRFDN 480
DB 421 EFVRLTNGHKGRIAGLOQYRDLVWSSSDNTIRLMDIEGACALVLEGHELVRCIRFDN 480
QY 481 KRIVSAGYDKIKWDLVAALDPRAPACTLCLRTLVHSGVFRLODFEFOIVSSHDDT 540
DB 481 KRIVSAGYDKIKWDLVAALDPRAPACTLCLRTLVHSGVFRLODFEFOIVSSHDDT 540
QY 541 ILIMDFLNDPAAOAEPPRSPRSTYTYISR 569
DB 541 ILIMDFLNDPAAOAEPPRSPRSTYTYISR 569

RESULT 9

US-09-700-444-12
; Sequence 12, Application US/09700444
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Marian R.
; TITLE OF INVENTION: CELL SIGNALING PROTEINS
; FILE REFERENCE: PF-0521 PCT
; CURRENT APPLICATION NUMBER: US/09/700,444
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/085,3434; 60/098,010
; PRIOR FILING DATE: 1998-05-13; 1998-08-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte Clone 3239149
US-09-700-444-12

Query Match 100.0%; Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPBAVLOEKALFKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
DB 1 MDPBAVLOEKALFKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
QY 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFFQWSESDQVEFEVHL 120
DB 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVYFFQWSESDQVEFEVHL 120
QY 121 ISOMCHYOHGHINSYLPKMLORDFITLAPARGDHIENLSTYLDKSLCAALVCKEY 180
DB 121 ISOMCHYOHGHINSYLPKMLORDFITLAPARGDHIENLSTYLDKSLCAALVCKEY 180
QY 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGQYLFKNRPPGNAPNSFYALYPKII 240
DB 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGQYLFKNRPPGNAPNSFYALYPKII 240
QY 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRNDWTIKIMDKNTLECK 300
DB 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRNDWTIKIMDKNTLECK 300
QY 301 RLUTGHTGSVLCLOQDEVIITGSSDSSTVRVWDVNTGEMLNTLLHHC EAVLHLRFNNGM 360
DB 301 RLUTGHTGSVLCLOQDEVIITGSSDSSTVRVWDVNTGEMLNTLLHHC EAVLHLRFNNGM 360
QY 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420
DB 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420

Db 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIYVSAGDRTIKYWNSTC 420
 Oy 421 EFWRLNKHKGIAICLOYRDRLVYSSSDNTIRLMDIEGACILRVLEGHEELVRCIRFDN 480
 Db 421 EFWRLNKHKGIAICLOYRDRLVYSSSDNTIRLMDIEGACILRVLEGHEELVRCIRFDN 480
 Oy 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540
 Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540
 Oy 541 ILIMDFLNDPAAQAEPSPRSRTTYTISR 569
 Db 541 ILIMDFLNDPAAQAEPSPRSRTTYTISR 569

RESULT 10

US-09-791-537-80169
 ; Sequence 80169, Application US/09/91537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: Patent version 3.0
 ; SEQ ID NO 80169
 ; LENGTH: 569
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-791-537-80169

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 1,3e-284;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MDPAEAVLDEKALKFNNSSEREDCNNGEPKRIPEKNSLRQYNSCARLCLNDQETVCLA 60
 Db 1 MDPAEAVLDEKALKFNNSSEREDCNNGEPKRIPEKNSLRQYNSCARLCLNDQETVCLA 60
 Oy 61 STAMKTEHCVAKTCLANGSSMIYVQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120
 Db 61 STAMKTEHCVAKTCLANGSSMIYVQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120
 Oy 121 ISOMCHYOHGHINSYLKPMLOPDTLALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180
 Db 121 ISOMCHYOHGHINSYLKPMLOPDTLALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180
 Oy 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERGMGOYLFRKPPGNAAPPNSFYALPKII 240
 Db 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERGMGOYLFRKPPGNAAPPNSFYALPKII 240
 Oy 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
 Db 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
 Oy 301 RILTGHTSVCLQYDEVEYITGSSDSTVAVWVNGEMNTLIIHCEAVLHLRFNNGM 360
 Db 301 RILTGHTSVCLQYDEVEYITGSSDSTVAVWVNGEMNTLIIHCEAVLHLRFNNGM 360
 Oy 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIYVSAGDRTIKYWNSTC 420
 Db 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIYVSAGDRTIKYWNSTC 420
 Oy 421 EFWRLNKHKGIAICLOYRDRLVYSSSDNTIRLMDIEGACILRVLEGHEELVRCIRFDN 480
 Db 421 EFWRLNKHKGIAICLOYRDRLVYSSSDNTIRLMDIEGACILRVLEGHEELVRCIRFDN 480
 Oy 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540
 Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540

Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540
 Oy 541 ILIMDFLNDPAAQAEPSPRSRTTYTISR 569
 Db 541 ILIMDFLNDPAAQAEPSPRSRTTYTISR 569

RESULT 11

US-09-832-161-18
 ; Sequence 18, Application US/09832161
 ; GENERAL INFORMATION:
 ; APPLICANT: Manning, Anthony M.
 ; APPLICANT: Mercurio, Frank
 ; APPLICANT: Amit, Sharon
 ; APPLICANT: Ben-Neriah, Yliron
 ; APPLICANT: Davis, Matti
 ; APPLICANT: Hatzuba, Ada
 ; APPLICANT: Lavon, Itis
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
 ; FILE REFERENCE: NE-KB
 ; CURRENT APPLICATION NUMBER: US/09/832,161
 ; CURRENT FILING DATE: 2001-04-09
 ; PRIOR FILING DATE: 1998-12-10
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 18
 ; LENGTH: 569
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-832-161-18

Query Match 100.0%; Score 3034; DB 22; Length 569;
 Best Local Similarity 100.0%; Pred. No. 1,3e-284;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MDPAEAVLDEKALKFNNSSEREDCNNGEPKRIPEKNSLRQYNSCARLCLNDQETVCLA 60
 Db 1 MDPAEAVLDEKALKFNNSSEREDCNNGEPKRIPEKNSLRQYNSCARLCLNDQETVCLA 60
 Oy 61 STAMKTEHCVAKTCLANGSSMIYVQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120
 Db 61 STAMKTEHCVAKTCLANGSSMIYVQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120
 Oy 121 ISOMCHYOHGHINSYLKPMLOPDTLALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180
 Db 121 ISOMCHYOHGHINSYLKPMLOPDTLALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180
 Oy 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERGMGOYLFRKPPGNAAPPNSFYALPKII 240
 Db 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERGMGOYLFRKPPGNAAPPNSFYALPKII 240
 Oy 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
 Db 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
 Oy 301 RILTGHTSVCLQYDEVEYITGSSDSTVAVWVNGEMNTLIIHCEAVLHLRFNNGM 360
 Db 301 RILTGHTSVCLQYDEVEYITGSSDSTVAVWVNGEMNTLIIHCEAVLHLRFNNGM 360
 Oy 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIYVSAGDRTIKYWNSTC 420
 Db 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIYVSAGDRTIKYWNSTC 420
 Oy 421 EFWRLNKHKGIAICLOYRDRLVYSSSDNTIRLMDIEGACILRVLEGHEELVRCIRFDN 480
 Db 421 EFWRLNKHKGIAICLOYRDRLVYSSSDNTIRLMDIEGACILRVLEGHEELVRCIRFDN 480
 Oy 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540
 Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540

QY 541 ILIWDPLNDPAQAEPSPRSTYTYISR 569
Db 541 ILIWDPLNDPAQAEPSPRSTYTYISR 569

RESULT 12

US-10-038-010-8
; Sequence 8, Application US/10038010
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4/67A
; CURRENT APPLICATION NUMBER: US/10/038,010
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: beta-TCP1
; LOCATION: (1). (569)
; OTHER INFORMATION:
US-10-038-010-8

Query Match 100.0%; Score 3034; DB 24; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLOEAKLKFMSNSEDNCNGEPKRIPEKNSLRQYNSCARLCLNGETVCLA 60
Db 1 MDPAAVLOEAKLKFMSNSEDNCNGEPKRIPEKNSLRQYNSCARLCLNGETVCLA 60
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOWSESQVEFEVHL 120
Db 61 STAMKTEVCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOWSESQVEFEVHL 120
QY 121 ISOMCHYGHINSYLPKMLQRFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
Db 121 ISOMCHYGHINSYLPKMLQRFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
QY 181 RYVSDGMLMKKLIERNVRTDSLRLGLAERRGWGYLFKNKPPDGNAPNSFYRALTYPKII 240
Db 181 RYVSDGMLMKKLIERNVRTDSLRLGLAERRGWGYLFKNKPPDGNAPNSFYRALTYPKII 240
QY 241 ODLETIESNMRGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
Db 241 ODLETIESNMRGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVLCLOYDERVITITGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILTGHTGSVLCLOYDERVITITGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRFNNGM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRLVYGHRAAVNVVDFDDKIVASASGDRITKYVNTSTC 420
Db 361 VTCSDKRSIAVWDMASPTDITLRLVYGHRAAVNVVDFDDKIVASASGDRITKYVNTSTC 420
QY 421 EFWRTLNHGRKGIACLOYRDLRVVSGSSDNTIRLMDIEGACLRVLEGHHEELVRCIRFDN 480
Db 421 EFWRTLNHGRKGIACLOYRDLRVVSGSSDNTIRLMDIEGACLRVLEGHHEELVRCIRFDN 480
QY 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540
Db 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540
QY 541 ILIWDPLNDPAQAEPSPRSTYTYISR 569
Db 541 ILIWDPLNDPAQAEPSPRSTYTYISR 569

RESULT 13
US-10-042-417-2
; Sequence 2, Application US/10042417
; GENERAL INFORMATION:
; APPLICANT: pagano, M.

; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-2

Query Match 100.0%; Score 3034; DB 24; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLOEAKLKFMSNSEDNCNGEPKRIPEKNSLRQYNSCARLCLNGETVCLA 60
Db 1 MDPAAVLOEAKLKFMSNSEDNCNGEPKRIPEKNSLRQYNSCARLCLNGETVCLA 60
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOWSESQVEFEVHL 120
Db 61 STAMKTEVCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOWSESQVEFEVHL 120
QY 121 ISOMCHYGHINSYLPKMLQRFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
Db 121 ISOMCHYGHINSYLPKMLQRFITLALPARGLDHAENILSYLDAKSLCAELVCKEMY 180
QY 181 RYVSDGMLMKKLIERNVRTDSLRLGLAERRGWGYLFKNKPPDGNAPNSFYRALTYPKII 240
Db 181 RYVSDGMLMKKLIERNVRTDSLRLGLAERRGWGYLFKNKPPDGNAPNSFYRALTYPKII 240
QY 241 ODLETIESNMRGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
Db 241 ODLETIESNMRGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVLCLOYDERVITITGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILTGHTGSVLCLOYDERVITITGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRFNNGM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRLVYGHRAAVNVVDFDDKIVASASGDRITKYVNTSTC 420
Db 361 VTCSDKRSIAVWDMASPTDITLRLVYGHRAAVNVVDFDDKIVASASGDRITKYVNTSTC 420
QY 421 EFWRTLNHGRKGIACLOYRDLRVVSGSSDNTIRLMDIEGACLRVLEGHHEELVRCIRFDN 480
Db 421 EFWRTLNHGRKGIACLOYRDLRVVSGSSDNTIRLMDIEGACLRVLEGHHEELVRCIRFDN 480
QY 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540
Db 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540
QY 541 ILIWDPLNDPAQAEPSPRSTYTYISR 569
Db 541 ILIWDPLNDPAQAEPSPRSTYTYISR 569

RESULT 14
US-60-098-010-5
; Sequence 5, Application US/60098010
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Tang, Y. Tom

```

; APPLICANT: Yang, Junming
; APPLICANT: Lal, Preeti
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Patterson, Chandra
; APPLICANT: Guejler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: G-PROTEIN SIGNALING PATHWAY MOLECULES
; FILE REFERENCE: PF-0583 P
; CURRENT APPLICATION NUMBER: US/60/098,010
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 569
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3239149
US-60-098-010-5

```

```

Query Match          100.0%; Score 3034; DB 27; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDPAEAVLQEKALKFPMNSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNOETVCLA 60
DB 1 MDPAEAVLQEKALKFPMNSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNOETVCLA 60
QY 61 STAKTEKNCVAKTKLANGTSSMIVPKOKLSASYEKKEKELCVKFEOMSESDQVEFEHL 120
DB 61 STAKTEKNCVAKTKLANGTSSMIVPKOKLSASYEKKEKELCVKFEOMSESDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
DB 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
QY 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
DB 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
QY 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
QY 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
QY 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
DB 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVLCLODYDERVITIGSSDSSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 RILTGHTGSVLCLODYDERVITIGSSDSSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
QY 361 VTCSKDRSIAVMDASPTDITLRVLVGHRAAVNVDPDDKIIVASGDRITIKVWNTSTC 420
DB 361 VTCSKDRSIAVMDASPTDITLRVLVGHRAAVNVDPDDKIIVASGDRITIKVWNTSTC 420
QY 421 EFVRTLNGHKRGIAVCLQYRDLVYVSGSSDNTIRLMDIEGACLVLEGHHELVRCIRFDN 480
DB 421 EFVRTLNGHKRGIAVCLQYRDLVYVSGSSDNTIRLMDIEGACLVLEGHHELVRCIRFDN 480
QY 481 KRIVSGAYDGIKIVMDLVVALDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
DB 481 KRIVSGAYDGIKIVMDLVVALDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
QY 541 ILIMDFLNDPAQAQAEPPRSPSRITYISR 569
DB 541 ILIMDFLNDPAQAQAEPPRSPSRITYISR 569

```

```

RESULT 15
US-60-098-355-2
; Sequence 2, Application US/60098355
; GENERAL INFORMATION:
; APPLICANT: Chaut, Dah Shiarn
; APPLICANT: Pagano, Michele
; APPLICANT: Latres, Esther
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS

```

```

; FILE REFERENCE: 5914-070
; CURRENT APPLICATION NUMBER: US/60/098,355
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PR
; ORGANISM: HUMAN
US-60-098-355-2

```

```

Query Match          100.0%; Score 3034; DB 27; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDPAEAVLQEKALKFPMNSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNOETVCLA 60
DB 1 MDPAEAVLQEKALKFPMNSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNOETVCLA 60
QY 61 STAKTEKNCVAKTKLANGTSSMIVPKOKLSASYEKKEKELCVKFEOMSESDQVEFEHL 120
DB 61 STAKTEKNCVAKTKLANGTSSMIVPKOKLSASYEKKEKELCVKFEOMSESDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
DB 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
QY 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
DB 121 ISOMCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEY 180
QY 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
QY 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLKKLIERVNRDTSIMRGIAERGGQYLFKKPPDGNAPPSFYRALYPKII 240
QY 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
DB 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVLCLODYDERVITIGSSDSSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 RILTGHTGSVLCLODYDERVITIGSSDSSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
QY 361 VTCSKDRSIAVMDASPTDITLRVLVGHRAAVNVDPDDKIIVASGDRITIKVWNTSTC 420
DB 361 VTCSKDRSIAVMDASPTDITLRVLVGHRAAVNVDPDDKIIVASGDRITIKVWNTSTC 420
QY 421 EFVRTLNGHKRGIAVCLQYRDLVYVSGSSDNTIRLMDIEGACLVLEGHHELVRCIRFDN 480
DB 421 EFVRTLNGHKRGIAVCLQYRDLVYVSGSSDNTIRLMDIEGACLVLEGHHELVRCIRFDN 480
QY 481 KRIVSGAYDGIKIVMDLVVALDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
DB 481 KRIVSGAYDGIKIVMDLVVALDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
QY 541 ILIMDFLNDPAQAQAEPPRSPSRITYISR 569
DB 541 ILIMDFLNDPAQAQAEPPRSPSRITYISR 569

```

```

Search completed: February 20, 2003, 09:58:44
Job time : 151 secs

```


FT Region 467..492
 FT /note="WD motif"
 FT 516..544
 FT /note="WD motif"
 XX W0938969-A1.
 XX
 XX 05-AUG-1999.
 XX
 XX 29-JAN-1999; 99WC-FR00196.
 XX
 XX 09-DEC-1998; 98FR-0015545.
 XX 30-JAN-1998; 98FR-0001100.
 XX
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX (INSP) INST PASTEUR.
 XX
 XX Arenzana Seisdedos F, Benarous R, Concordet J, Durand H;
 XX Kroll M, Margottin F;
 XX WPI: 1999-469329/39.
 XX N-PSDB; AAX86501.
 XX
 XX New human beta-transducin repeat containing protein and its
 XX fragments useful as, or to screen for, antiviral, antitumour,
 XX anti-inflammatory and anti-Alzheimer's agents
 XX
 XX Claim 1: Page 60-61; 71pp; French.
 XX
 XX The present sequence represents a human beta-transducin repeat containing
 XX protein (beta-Trcp). The protein directs proteins to the proteosome
 XX degradation pathways. The protein is able to interact with the Vpu
 XX protein of human immune deficiency virus-1 (HIV-1), cellular proteins
 XX IkappaB or beta-catenin (bc) and/or protein Skp1. The protein controls
 XX ubiquitinylation of phosphorylated proteins and thus their targeting to
 XX proteosomes for degradation. Depending on whether the process is
 XX inhibited or promoted, the result may be delayed breakdown of CD4 (in
 XX cases of HIV-1 infection); increased activity of Ikb (and thus reduced
 XX activity of NFkappaB); increased degradation of mutant bc in tumour
 XX cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
 XX patients. The beta-Trcp protein, and its active peptide fragments, or its
 XX nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
 XX antitumour agents that disrupt cell cycle regulation or protein
 XX degradation in human tumour cells, and anti-inflammatory agents that
 XX disrupt activation by NFkappaB. Fragments of the protein are also
 XX useful for treating osteo-articular inflammation or acute inflammation
 XX associated with release of tumour necrosis factor.
 XX
 XX Sequence 569 AA;
 XX
 XX Query Match 100.0%; Score 3034; DB 20; Length 569;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e-286;
 XX Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 MDPAAVAVLQEKALFKMNSREDNCGNPPKRIPEKNSLRQVNSCARCLINQETVCLA 60
 XX 1 MDPAAVAVLQEKALFKMNSREDNCGNPPKRIPEKNSLRQVNSCARCLINQETVCLA 60
 XX
 XX 61 STAMTENCVAKTLANGTSSMIVPKOKKLSASTEKEKELCVKFEQMSSEDOVEFEHL 120
 XX 61 STAMTENCVAKTLANGTSSMIVPKOKKLSASTEKEKELCVKFEQMSSEDOVEFEHL 120
 XX
 XX 121 ISOMCHYHGHINSYLRQDFITLALPARGLDHINILSTYDAKSLCAAEVYCKEWY 180
 XX 121 ISOMCHYHGHINSYLRQDFITLALPARGLDHINILSTYDAKSLCAAEVYCKEWY 180
 XX
 XX 121 ISOMCHYHGHINSYLRQDFITLALPARGLDHINILSTYDAKSLCAAEVYCKEWY 180
 XX 121 ISOMCHYHGHINSYLRQDFITLALPARGLDHINILSTYDAKSLCAAEVYCKEWY 180
 XX
 XX 181 RYTSQGMIMKKLIERMYVTDLSLMLGLAERRGCGYLFEKNPPQGNAPNPFYALPKII 240
 XX 181 RYTSQGMIMKKLIERMYVTDLSLMLGLAERRGCGYLFEKNPPQGNAPNPFYALPKII 240
 XX
 XX 181 RYTSQGMIMKKLIERMYVTDLSLMLGLAERRGCGYLFEKNPPQGNAPNPFYALPKII 240
 XX 181 RYTSQGMIMKKLIERMYVTDLSLMLGLAERRGCGYLFEKNPPQGNAPNPFYALPKII 240
 XX
 XX 241 ODIEETESNMRCGRHSIORITCRSETSKGYCLOYDQKIVSGLRDNTIKIMDKNTLECK 300
 XX 241 ODIEETESNMRCGRHSIORITCRSETSKGYCLOYDQKIVSGLRDNTIKIMDKNTLECK 300

QY 301 RILTHGHSVLCLOYDERVITITGSSDSTVAVMPOVNTGEMNLTIHHCFAVLHFRNGM 360
 DB 301 RILTHGHSVLCLOYDERVITITGSSDSTVAVMPOVNTGEMNLTIHHCFAVLHFRNGM 360
 QY 361 VTCSKDRSIAMWMASTPTDITLRRVLVGRRAAVNVDPDKYIVSASGRTIKVNTSTC 420
 DB 361 VTCSKDRSIAMWMASTPTDITLRRVLVGRRAAVNVDPDKYIVSASGRTIKVNTSTC 420
 QY 421 EFVRLTHGHSVLCLOYDERVITITGSSDSTVAVMPOVNTGEMNLTIHHCFAVLHFRNGM 480
 DB 421 EFVRLTHGHSVLCLOYDERVITITGSSDSTVAVMPOVNTGEMNLTIHHCFAVLHFRNGM 480
 QY 481 KRIYSGAYDGKIKVMDLVAAIDPRAPAGTLCRTLVHSGRAVRLQDFEFOVSSHDT 540
 DB 481 KRIYSGAYDGKIKVMDLVAAIDPRAPAGTLCRTLVHSGRAVRLQDFEFOVSSHDT 540
 QY 541 ILIMDFLNDPAAQAEPPRSPRTYTYISR 569
 DB 541 ILIMDFLNDPAAQAEPPRSPRTYTYISR 569

RESULT 2
 AAB12813
 ID AAB12813 standard; protein; 569 AA.

XX AAB12813;
 XX
 XX 27-NOV-2000 (first entry)
 XX
 XX Human beta-transducin repeat containing protein (beta-Trcp) S60 ID NO:3.
 XX
 XX Ubiquitin ligase SCF complex; F-box protein; ubiquitinylation; IkappaB;
 XX beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;
 XX gene therapy; colon cancer; beta-transducin repeat containing protein;
 XX beta-Trcp.
 XX
 XX Homo sapiens.
 XX
 XX JP2000166542-A.
 XX
 XX 20-JUN-2000.
 XX
 XX 02-DEC-1998; 98JP-0343437.
 XX
 XX 02-DEC-1998; 98JP-0343437.
 XX
 XX (KAGA-) KAGAKU GIUTUSU SHINKO JIGYODAN.
 XX
 XX WPI: 2000-485550/43.
 XX N-PSDB; AAA73132.
 XX
 XX F-box protein of ubiquitin ligase SCF complex which promotes the
 XX ubiquitinylation of IkappaB or beta-catenin
 XX
 XX Claim 3; Page 10-12; 19pp; Japanese.

XX The present invention describes an F-box motif protein of ubiquitin
 XX ligase SCF complex which promotes the ubiquitinylation of IkappaB or
 XX beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
 XX complex (SCF complex) of F-box protein containing F-box motif and WD40
 XX repeat motif and has the amino acid sequence of 45 residues (AAB12813)
 XX or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
 XX ligase FMD1 protein) and (AAB12813, which is human beta-transducin
 XX repeat containing protein (beta-Trcp)). The F-box protein can be used for
 XX the gene therapy of colon cancer by being recombined to a virus vector.
 XX
 XX Sequence 569 AA;
 XX

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 1.4e-286;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVIOEKALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60
 DB 1 MDPAAVIOEKALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60
 QY 61 STAMKTEKNCVAKTKLANGSSMIVPKORRLASYEKEKELCYKFEQMSDOVEFEVHL 120
 DB 61 STAMKTEKNCVAKTKLANGSSMIVPKORRLASYEKEKELCYKFEQMSDOVEFEVHL 120
 QY 121 ISOMCHYOCHGHIINSYLKPMLODFITLALPARGLDHIHAENILSYLDAKSLCAAEIVCKEMY 180
 DB 121 ISOMCHYOCHGHIINSYLKPMLODFITLALPARGLDHIHAENILSYLDAKSLCAAEIVCKEMY 180
 QY 181 RYTSOGLMKKLIERNVTRDLSMRGLAERRGQGYLFKNKPPDGNAPNSFYRALYPKII 240
 DB 181 RYTSOGLMKKLIERNVTRDLSMRGLAERRGQGYLFKNKPPDGNAPNSFYRALYPKII 240
 QY 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGSVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 RILTGHTGSVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 QY 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVVDDDKYIVSASGDRITKWNSTJC 420
 DB 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVVDDDKYIVSASGDRITKWNSTJC 420
 QY 421 EFVRTLNGHKGRIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVRIRFND 480
 DB 421 EFVRTLNGHKGRIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVRIRFND 480
 QY 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 DB 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 QY 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569
 DB 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569

RESULT 3
 AAY96697
 ID AAY96697 standard; Protein; 569 AA.
 AC AAY96697;
 XX 26-SEP-2000 (first entry)
 DT 26-SEP-2000 (first entry)
 DE Human beta-TrCP.
 KM E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
 KM nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
 KM anti-inflammatory; immunosuppressive; cytoskeletal.
 OS Homo sapiens.
 XX MO200034447-A2.
 XX 15-JUN-2000.
 PF 10-DEC-1999; 99MO-US29371.
 PR 10-DEC-1998; 98US-0210060.
 XX (SIGN-) SIGNAL PHARM INC.
 PA (YISS) YISSUM RES & DEV CO.
 PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
 PI Lavan I, Yaron A;
 XX WPI: 2000-431294/37.
 DR N-PSDB; AAS1229.

XX poly peptide enhancing phosphorylated I-kappa-B ubiquitination useful for
 PT treating disorder associated with NF-kappa-B activation e.g. cancer,
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its
 PT variant
 PS Claim 21; Page 72-74; 77pp; English.
 XX Human beta-TrCP, an F-box/WD protein family member, has been shown to
 CC have homology to human E3 ubiquitin ligase (E3). E3 enhances
 CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of
 CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
 CC degradation via the ubiquitin pathway is useful for identifying
 CC modulators of this process for use in treating diseases associated with
 CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
 CC the F-box results in a protein that functions as a dominant negative
 CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a
 CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
 CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
 CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
 CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
 CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
 XX Sequence 569 AA:
 SQ
 Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 1,4e-286;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDPAAVIOEKALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60
 DB 1 MDPAAVIOEKALKFNMSSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60
 QY 61 STAMKTEKNCVAKTKLANGSSMIVPKORRLASYEKEKELCYKFEQMSDOVEFEVHL 120
 DB 61 STAMKTEKNCVAKTKLANGSSMIVPKORRLASYEKEKELCYKFEQMSDOVEFEVHL 120
 QY 121 ISOMCHYOCHGHIINSYLKPMLODFITLALPARGLDHIHAENILSYLDAKSLCAAEIVCKEMY 180
 DB 121 ISOMCHYOCHGHIINSYLKPMLODFITLALPARGLDHIHAENILSYLDAKSLCAAEIVCKEMY 180
 QY 181 RYTSOGLMKKLIERNVTRDLSMRGLAERRGQGYLFKNKPPDGNAPNSFYRALYPKII 240
 DB 181 RYTSOGLMKKLIERNVTRDLSMRGLAERRGQGYLFKNKPPDGNAPNSFYRALYPKII 240
 QY 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGSVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 RILTGHTGSVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 QY 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVVDDDKYIVSASGDRITKWNSTJC 420
 DB 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVVDDDKYIVSASGDRITKWNSTJC 420
 QY 421 EFVRTLNGHKGRIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVRIRFND 480
 DB 421 EFVRTLNGHKGRIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEGEHELVRIRFND 480
 QY 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 DB 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 QY 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569
 DB 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569

RESULT 4
 AAY83041
 ID AAY83041 standard; Protein; 569 AA.

XX AC AAY83041;
 XX DX 16-AUG-2000 (first entry)
 XX DE F-box protein FBP-1.
 XX F-box protein; FBP; diagnosis; treatment; screening; agonist;
 KW antagonist; proliferative disorder; differentiative disorder;
 KW breast cancer; prostate cancer; ovarian cancer; cancer;
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
 KW inflammatory disorder; human.

XX OS Homo sapiens.
 XX PN MO200012679-A1.

XX PD 09-MAR-2000.

XX PF 27-AUG-1999; 99WO-US19560.

XX PR 28-AUG-1998; 98US-0098355.

XX PR 03-FEB-1999; 99US-0118568.

XX PR 15-MAR-1999; 99US-0124449.

XX PA (UYNY) UNIV NEW YORK STATE.

XX PI Chlaur DS, Pagano M, Latres E;
 XX DX WPI: 2000-256635/22.
 XX DR N-PSDB; AA293350.

XX PT Novel nucleic acid for screening compounds useful for treating
 XX PT proliferative and differentiative disorders such as cancer and immune
 XX PT disorders comprises sequences encoding ubiquitin ligases
 XX PS Disclosure; Figure 3a; 245pp; English.

XX CC Nucleic acids encoding substrate-targeting subunits of ubiquitin
 XX CC ligases with F-box motifs (F-box proteins) are useful for diagnosis
 XX CC of proliferative and differentiated related disorders by measuring
 XX CC FBP gene expression. Cells expressing such proteins or
 XX CC their fragments are useful for screening compounds. The compounds
 XX CC are agonists or antagonists, which are useful for treating a
 XX CC proliferative or differentiative disorder in a mammal such as
 XX CC breast, ovarian and prostate cancer and small cell lung carcinoma
 XX CC and also major opportunistic infections, immune disorders,
 XX CC cardiovascular diseases and inflammatory disorders. FBP protein,
 XX CC analogs, derivatives and their subsequences, anti-FBP antibodies
 XX CC are also useful in diagnosis of the disorders.

XX SQ Sequence 569 AA;

Query Match 100.0%; Score 3034; DB 21; Length 569;

Best Local Similarity 100.0%; Pred. No. 1.4e-286; Mismatches 0; Indels 0; Gaps 0;

XX 1 MDPAAVLTQKALKFNNSSREDCNNGEPPKIIPEKNSLRQYNSCARCLNQEIVCIA 60
 XX DB 1 MDPAAVLTQKALKFNNSSREDCNNGEPPKIIPEKNSLRQYNSCARCLNQEIVCIA 60
 XX QY 61 STAMKTENCYAKTKLANGISSMTVPKORKLSASYEKEKELCVYFEQWESDQVEFEHL 120
 XX DB 61 STAMKTENCYAKTKLANGISSMTVPKORKLSASYEKEKELCVYFEQWESDQVEFEHL 120
 XX QY 121 ISOMCHVGHGHSYKPMORDFITAPRGIDHTAENITSLDIAKSKCAAEVLCKEY 180
 XX DB 121 ISOMCHVGHGHSYKPMORDFITAPRGIDHTAENITSLDIAKSKCAAEVLCKEY 180
 XX QY 181 RYVSDGMLMKLLIERVYRTDLSLRGLAERKQGYLFKNKPPDGNAPPSFYRATVPKII 240
 XX DB 181 RYVSDGMLMKLLIERVYRTDLSLRGLAERKQGYLFKNKPPDGNAPPSFYRATVPKII 240

QY 241 QDIETIESMWRGHSLSORICHSSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
 DB 241 QDIETIESMWRGHSLSORICHSSETSKGVYCLQYDDOKIVSGLRDNITIKIMDKNTLECK 300
 QY 301 RILTGHGTSYLCQYDERVITIGSSDSYVAVMDVNTGEMLNLIIHCEAVLHRRNNGM 360
 DB 301 RILTGHGTSYLCQYDERVITIGSSDSYVAVMDVNTGEMLNLIIHCEAVLHRRNNGM 360
 QY 361 VYCSKRSIAVMDASPTITLRRVLVGHRAAVNVDPDKIVASGDRITKWNSTSC 420
 DB 361 VYCSKRSIAVMDASPTITLRRVLVGHRAAVNVDPDKIVASGDRITKWNSTSC 420
 QY 421 EYVITLNGHKGITACIQYDRDLVSSGSDNTIRLMDIEGACRLVLEHLELVCIRFDN 480
 DB 421 EYVITLNGHKGITACIQYDRDLVSSGSDNTIRLMDIEGACRLVLEHLELVCIRFDN 480
 QY 481 KRIVSGAYDGKITKWDVLVALDPRAPAGTLCRTLVESHGVRFLQDFEFQIVSSHDDT 540
 DB 481 KRIVSGAYDGKITKWDVLVALDPRAPAGTLCRTLVESHGVRFLQDFEFQIVSSHDDT 540
 QY 541 IILWDFLNDPAAQAEPPRSPSRITYISR 569
 DB 541 IILWDFLNDPAAQAEPPRSPSRITYISR 569

RESULT 5

AA83250
 ID AAY83250 standard; Protein; 569 AA.

XX AC AAY83250;

XX DX 16-AUG-2000 (first entry)

XX DE F-box protein hbetaTrCp.

XX KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 XX KW destabilization; proteolysis; drug discovery; gene therapy; cancer;
 XX KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
 XX KW human.

XX OS Homo sapiens.

XX PN WO200022110-A2.

XX PD 20-APR-2000.

XX PF 08-OCT-1999; 99WO-US23705.

XX PR 09-OCT-1998; 98US-0103787.

XX PA (HARD) HARVARD COLLEGE.

XX PI Zhou P, Howley P;

XX DX WPI: 2000-317970/27.

XX DR N-PSDB; AA293710.

XX PT Targeting degradation of polypeptide useful for treating cancer and
 XX PT other proliferative disorders. Involves conjugating polypeptide with
 XX PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 XX PT compound

XX PS Claim 9; Page 171; 185pp; English.

XX CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 XX CC ligases) which can be used for the targeted degradation of a target
 XX CC polypeptide in vivo. Targeted degradation is achieved by expressing
 XX CC the ubiquitin ligase in a cell linked to the interaction domain of
 XX CC the target polypeptide and thereby recruiting the target polypeptide
 XX CC to the ubiquitin ligase. Such methods are useful for decreasing or
 XX CC increasing the level of a target polypeptide and for creating and
 XX CC expressing a destabilized polypeptide which is subjected to SCF
 XX CC mediated proteolysis. Degrading any desired protein in a cell is

PS Claim 1; Page 77-78; 90pp; English.
 CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded
 CC by cDNA obtained from Inocyte clone 3239149 of COLAD0701 library. It is
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is
 CC found to be homologous to beta-transducin repeats containing
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as
 CC hybridisation probe for detecting CSIGP related sequences or allelic
 CC variants. Recombinant CSIGP can be produced in host cells by transforming
 CC them with genetically engineered vectors. Agonists or antagonists can be
 CC used in the treatment of cell proliferative and inflammatory disorders
 CC associated with decreased or increased CSIGP expression. CSIGP is used in
 CC the diagnosis, prevention and treatment of cell proliferative disorders
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.

Sequence 569 AA:

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 1,4e-286;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLOEKAALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARCLNQEYVCLA 60
 DB 1 MDPAAVLOEKAALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARCLNQEYVCLA 60

QY 61 STAMKTEKNCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQWSSDQVEVEHL 120
 DB 61 STAMKTEKNCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQWSSDQVEVEHL 120

QY 121 ISOMCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180
 DB 121 ISOMCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180

QY 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERRGQGYLFKNKPPGNAFPNSFYRALYPKII 240
 DB 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERRGQGYLFKNKPPGNAFPNSFYRALYPKII 240

QY 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMKNTLECK 300
 DB 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMKNTLECK 300

QY 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDYNTGEMLNTLIHCEAVLHLRFNNGM 360
 DB 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDYNTGEMLNTLIHCEAVLHLRFNNGM 360

QY 361 VTCSKDRSIAVWMAASPTDITLRRLVGHRAAVNVDFDDKYIVSASGDRITIKVWNTSTC 420
 DB 361 VTCSKDRSIAVWMAASPTDITLRRLVGHRAAVNVDFDDKYIVSASGDRITIKVWNTSTC 420

QY 421 EFVRTLNGHKGRIACLOYRDRLVVGSSDNTIRLMDIEGACLRVEGHEELVRCIRFPN 480
 DB 421 EFVRTLNGHKGRIACLOYRDRLVVGSSDNTIRLMDIEGACLRVEGHEELVRCIRFPN 480

QY 481 KRIVSAGYOKITVMDLVVALDPRAGTLCRTLVESHGGRVFRLODFEQIVSSSHDT 540
 DB 481 KRIVSAGYOKITVMDLVVALDPRAGTLCRTLVESHGGRVFRLODFEQIVSSSHDT 540

QY 541 ILITMDFLNDPAAQAEPPRSPRTYTYISR 569
 DB 541 ILITMDFLNDPAAQAEPPRSPRTYTYISR 569

RESULT 7
 AAB48298
 ID AAB48298 standard; protein; 569 AA.

AC AAB48298;

DT 02-APR-2001 (first entry)

DE Human ZP11 protein.

XX

KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
 KW Bad; Bcl-2; tumour; cytosolic.

OS Homo sapiens.

PN WO200075184-A1.

PD 14-DEC-2000.

PF 05-JUN-2000; 2000WO-US15449.

PR 04-JUN-1999; 99US-0137494.

PA (UYUA) UNIV YALE.

PI Zhang H, Tsvetkov IM, Kondo T;

DR WPI: 2001-061703/07.

DR N-PSDB: AAC84610.

PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
 PT involves altering levels of proteins such as S-phase kinase associated
 PT proteins 1, 2 and cullin/CDC53 proteins -

PS Claim 3; Page 130-132; 162pp; English.

The invention relates to methods of altering the polypeptide levels in a
 cell, using proteins selected from S-phase kinase associated proteins 1
 and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
 cullin/CDC53 family of proteins). The method is useful for altering the
 level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
 polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
 detecting tumours, and in monitoring tumor treatment in a mammal. Agents
 that modulate interactions between SKP and target proteins are useful for
 treating tumours.

Sequence 569 AA:

Query Match 99.8%; Score 3027; DB 22; Length 569;
 Best Local Similarity 99.8%; Pred. No. 6,9e-286;
 Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDPAAVLOEKAALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARCLNQEYVCLA 60
 DB 1 MDPAAVLOEKAALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARCLNQEYVCLA 60

QY 61 STAMKTEKNCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQWSSDQVEVEHL 120
 DB 61 STAMKTEKNCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQWSSDQVEVEHL 120

QY 121 ISOMCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180
 DB 121 ISOMCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180

QY 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERRGQGYLFKNKPPGNAFPNSFYRALYPKII 240
 DB 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERRGQGYLFKNKPPGNAFPNSFYRALYPKII 240

QY 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMKNTLECK 300
 DB 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMKNTLECK 300

QY 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDYNTGEMLNTLIHCEAVLHLRFNNGM 360
 DB 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDYNTGEMLNTLIHCEAVLHLRFNNGM 360

QY 361 VTCSKDRSIAVWMAASPTDITLRRLVGHRAAVNVDFDDKYIVSASGDRITIKVWNTSTC 420
 DB 361 VTCSKDRSIAVWMAASPTDITLRRLVGHRAAVNVDFDDKYIVSASGDRITIKVWNTSTC 420

QY 421 EFVRTLNGHKGRIACLOYRDRLVVGSSDNTIRLMDIEGACLRVEGHEELVRCIRFPN 480
 DB 421 EFVRTLNGHKGRIACLOYRDRLVVGSSDNTIRLMDIEGACLRVEGHEELVRCIRFPN 480

XX

DB 421 EFVRLNGHKGKGIACIQYRDRLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
 QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQFDEFOIVSSSHDT 540
 DB 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQFDEFOIVSSSHDT 540
 QY 541 ILIWDPLNDPAAQAEPPRSPRTTYTISR 569
 DB 541 ILIWDPLNDPAAQAEPPRSPRTTYTISR 569
 RESULT 8
 AAM00960
 ID AAM00960 standard; Protein: 608 AA.
 AC AAM00960;
 XX 01-OCT-2001 (first entry)
 XX Human bone marrow protein, SEQ ID NO: 436.
 DE Human: bone marrow; antinflammatory; cytostatic; neuroprotective;
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KW immunosuppressive; gene therapy; cytokine cell proliferation;
 KW cell differentiation modulator; immune disorder; infection; cancer;
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
 XX Homo sapiens.
 OS
 PN W0200153453-A2.
 PD 26-JUL-2001.
 XX 23-DEC-2000; 2000WO-US34960.
 PE 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX (HYSEQ-) HYSEQ INC.
 PA Ford JB, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 XX WPI; 2001-488707/53.
 DR N-PSDB; AAH90079.
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -
 PS Claim 10; Page 523-524; 648pp; English.
 CC The present sequence is one of 251 novel human polynucleotides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.

SQ Sequence 608 AA:
 Query Match 99.8%; Score 3027; DB 22; Length 608;
 Best Local Similarity 99.6%; Pred. No. 7.7e-286;
 Matches 567; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDPAEAVIOEAKLKFNSSEREDCNGEPKRIPEKNSLRQTSYSCARLCLNDVCL 60
 DB 40 MDPEAVIOEAKLKFNSSEREDCNGEPKRIPEKNSLRQTSYSCARLCLNDVCL 99
 QY 61 STAMKTEHCVAATKLANGSSMTIVPRORLASYSYKEKELCYKFEOMSESQVFEVHL 120
 DB 100 STAMKTEHCVAATKLANGSSMTIVPRORLASYSYKEKELCYKFEOMSESQVFEVHL 159
 QY 121 ISOMCHYOHGHINSYLPMLQDRFTALPARGLDHTAENILSYLDKSLCAAEVCKEM 180
 DB 160 ISOMCHYOHGHINSYLPMLQDRFTALPARGLDHTAENILSYLDKSLCAAEVCKEM 219
 QY 181 RYTSDEMLMKKLIERNVPTDSLIRGLAERRGQYLFKKPPDGNAPPSFYRALYPKII 240
 DB 220 RYTSDEMLMKKLIERNVPTDSLIRGLAERRGQYLFKKPPDGNAPPSFYRALYPKII 279
 QY 241 QDIETTESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
 DB 280 QDIETTESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 339
 QY 301 RILGHTGSVYLCQYDERVYITGSSDSSTVYVMDVNTGEMNLIRHCEAVLHLPFNCGM 360
 DB 340 RILGHTGSVYLCQYDERVYITGSSDSSTVYVMDVNTGEMNLIRHCEAVLHLPFNCGM 399
 QY 361 VTCSKDRSTAVWDMASPTDITLRLVYVGHRAVNVYDFDDKTYVSGSRTIKVWNTSYC 420
 DB 400 VTCSKDRSTAVWDMASPTDITLRLVYVGHRAVNVYDFDDKTYVSGSRTIKVWNTSYC 459
 QY 421 EFVRLNGHKGKGIACIQYRDRLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
 DB 460 EFVRLNGHKGKGIACIQYRDRLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 519
 QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQFDEFOIVSSSHDT 540
 DB 520 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQFDEFOIVSSSHDT 579
 QY 541 ILIWDPLNDPAAQAEPPRSPRTTYTISR 569
 DB 580 ILIWDPLNDPAAQAEPPRSPRTTYTISR 608
 RESULT 9
 AAM78582
 ID AAM78582 standard; Protein: 605 AA.
 AC AAM78582;
 XX 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 1244.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 PN W0200157190-A2.
 PD 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US04098.
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663361.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR N-PSDB: AAK51715.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PS Claim 20; Page 3503-3504; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAH80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibit activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX Sequence 605 AA:
 SQ
 Query Match 99.1%; Score 3006; DB 22; Length 605;
 Best Local Similarity 94.0%; Pred. No. 8.5e-284;
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;
 QY 1 MDPAEAVLQEKALK-----FNNSSREDC 24
 DB 1 MDPAEAVLQEKALKFNNSSREDC 60
 QY 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSSMTV 84
 DB 61 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSSMTV 120
 QY 85 PRORIKASAYEKEKELCVYFFQWSESDOYEFVEHLISQCHYOHGHNISYLKPMLODF 144
 DB 121 PRORIKASAYEKEKELCVYFFQWSESDOYEFVEHLISQCHYOHGHNISYLKPMLODF 180
 QY 145 ITALPARGLDHAENILSYLDKASLCALVELCKEYRYATSDMKLKKLIERVRRDSIMR 204
 DB 181 ITALPARGLDHAENILSYLDKASLCALVELCKEYRYATSDMKLKKLIERVRRDSIMR 240
 QY 205 GLAERGMQCYLFKNKPPDGNAPNSFYRALPKIIODIETIESNMGCRHSIQRIHRS 264
 DB 241 GLAERGMQCYLFKNKPPDGNAPNSFYRALPKIIODIETIESNMGCRHSIQRIHRS 300
 QY 265 ETSKGVYCLQYDDOKIVSLGRNTIKIMDKNTLECKRILITGHTGVLQYDERVITIGS 324
 DB 301 ETSKGVYCLQYDDOKIVSLGRNTIKIMDKNTLECKRILITGHTGVLQYDERVITIGS 360
 QY 325 SDSTVAVDVNTGEMINLTIHCEAVLHLRFNNGMAYTCSKDRSTAVYDMSPTDITLR 384
 DB 361 SDSTVAVDVNTGEMINLTIHCEAVLHLRFNNGMAYTCSKDRSTAVYDMSPTDITLR 420
 QY 385 VLVGRRAAVNVDPDDKTIYASGDRITIKVNTSTCEFEVRLTNGHKGSIACIQYRDLRV 444
 DB 421 VLVGRRAAVNVDPDDKTIYASGDRITIKVNTSTCEFEVRLTNGHKGSIACIQYRDLRV 480

QY 445 SGSSDNTIRLMDIEGACLRVLEGHEILVRCIRFDRNRIVSGAYDGKIKWMDLVAALDP 504
 DB 481 SGSSDNTIRLMDIEGACLRVLEGHEILVRCIRFDRNRIVSGAYDGKIKWMDLVAALDP 540
 QY 505 APAGTLCRLTVLHSGRVRLQDFEFQIVSSSHDPTILMDLNDPAAQAEPPSPSRXY 564
 DB 541 APAGTLCRLTVLHSGRVRLQDFEFQIVSSSHDPTILMDLNDPAAQAEPPSPSRXY 600
 QY 565 TYISR 569
 DB 601 TYISR 605
 RESULT 10
 AAB12812
 ID AAB12812 standard; protein; 569 AA.
 AC AAB12812;
 DT 27-NOV-2000 (first entry)
 DE Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.
 KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;
 KW beta-catenin; Skp1; Cull; F-box motif; WD40 repeat motif; FWD1;
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;
 KW beta-TrCP.
 OS Mus musculus.
 PN JP2000166542-A.
 PD 20-JUN-2000.
 XX 02-DEC-1998; 98JP-0343437.
 XX 02-DEC-1998; 98JP-0343437.
 PR (KAGA-) KAGAU GIJUTSU SHINKO JIGYODAN.
 PA WPI: 2000-485550/43.
 DR N-PSDB: AAA73131.
 DR F-box protein of ubiquitin ligase SCF complex which promotes the
 PT ubiquitination of IkappaB or beta-catenin -
 PT ubiquitination of IkappaB or beta-catenin -
 PS Claim 2; Page 9-10; 19pp; Japanese.
 CC The present invention describes an F-box motif protein of ubiquitin
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or
 CC beta-catenin and is constituted by Skp1 protein, Cull protein and a
 CC complex (SCF complex) of F-box protein containing F-box motif and a
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
 CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for
 CC the gene therapy of colon cancer by being recombined to a virus vector.
 CC
 XX Sequence 569 AA:
 SQ
 Query Match 98.8%; Score 2997; DB 21; Length 569;
 Best Local Similarity 98.6%; Pred. No. 5.8e-283;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MDPAEAVLQEKALKFNNSSREDCNNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCL 60
 DB 1 MDPAEAVLQEKALKFNNSSREDCNNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCL 60
 QY 61 STAMKTENCYAKTKLANGTSSMTVPRORIKASAYEKEKELCVYFFQWSESDOYEFVEHL 120
 DB 61 STAMKTENCYAKTKLANGTSSMTVPRORIKASAYEKEKELCVYFFQWSESDOYEFVEHL 120
 QY 121 ISQCHYOHGHNISYLKPMLODFITALPARGLDHAENILSYLDKASLCALVELCKEY 180

```

Db 121 ISQCHYOHGHINSYLPKMLQDRDTITLPAKGLDHIENLITSLDAKSLCAAEIVCKEY 180
OY 181 RYISDGLMKKLLERWRTDSLMWGLAERGMGOYLFKNPPDGNAPNSFYALPKII 240
Db 181 RYISDGLMKKLLERWRTDSLMWGLAERGMGOYLFKNPPDGNAPNSFYALPKII 240
OY 241 ODITISNMRCGRHSIQRHCHSETSKGYCLODYDOKIVSGLRDNTIKIMDKNTLECK 300
Db 241 ODITISNMRCGRHSIQRHCHSETSKGYCLODYDOKIVSGLRDNTIKIMDKNTLECK 300
OY 301 RILGHTGSVLCLODYDERVYITGSSDSTVAVWVNTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILGHTGSVLCLODYDERVYITGSSDSTVAVWVNTGEMLNTLIHCEAVLHLRFNNGM 360
OY 361 VTCKSDRSIAVWMASTPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
Db 361 VTCKSDRSIAVWMASTPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
OY 421 EFVRTLNGHKGIGACLOYRDRLVYVSSSDNTIRLMDIECGACLVLEGHELVRCIRFDN 480
Db 421 EFVRTLNGHKGIGACLOYRDRLVYVSSSDNTIRLMDIECGACLVLEGHELVRCIRFDN 480
OY 481 KRIVSGAYDGKIKYWDLVAAIDPPAPAGTLCRLTVLHSGRVFLQDFEOLVSSSHDT 540
Db 481 KRIVSGAYDGKIKYWDLVAAIDPPAPAGTLCRLTVLHSGRVFLQDFEOLVSSSHDT 540
OY 541 ILIMDFLNDPAAQAEPSPSRITYTISR 569
Db 541 ILIMDFLNDPAAQAEPSPSRITYTISR 569

```

RESULT 11
AA83254
ID AAY83254 standard; Protein: 569 AA.

```

AC AAY83254;
XX 16-AUG-2000 (first entry)
DT F-box protein FMDlp.
XX
XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;
XX destabilization; proteolysis; drug discovery; gene therapy; cancer;
XX oncoprotein; Huntington's disease; gene knockout; delivery systems;
XX mouse; ss.
XX Mus musculus.
XX OS
XX WO200022110-A2.
XX 20-APR-2000.
XX 08-OCT-1999; 99WO-US23705.
XX 09-OCT-1998; 98US-0103787.
XX (HARD) HARVARD COLLEGE.
XX Zhou P, Howley P;
XX WPI; 2000-317970/27.
XX N-PSDB; AA293714.
XX Targeting degradation of polypeptide useful for treating cancer and
XX other proliferative disorders; involves conjugating polypeptide with
XX ubiquitin protein ligase or inhibiting ubiquitination using organic
XX compound
XX Claim 9; Page 184-185; 185pp; English.
XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
XX ligases) which can be used for the targeted degradation of a target

```

polypeptide in vivo. Targeted degradation is achieved by expressing the ubiquitin ligase in a cell linked to the interaction domain of the target polypeptide and thereby recruiting the target polypeptide to the ubiquitin ligase. Such methods are useful for decreasing or increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating diseases caused by the presence of abnormal amount of the specific polypeptides, for drug discovery and for gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

Sequence 569 AA:

Query Match 98.8%; Score 2997; DB 21; Length 569;
Best local Similarity 98.6%; Pred. No. 5,8e-283;
Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 MDPAEAVLOEKALKFNNSSREDCCNGEPPRKIIPEKNSLRQYNSCARLCLNOETVCLA 60
Db 1 MDPAEAVLOEKALKFNNSSREDCCNGEPPRKIIPEKNSLRQYNSCARLCLNOETVCLA 60
OY 61 STAMKTEVCYAKTKLANGTSSMIVPQKRLSASEKEKELCYVFEQWESDQVEVEHL 120
Db 61 STAMKTEVCYAKTKLANGTSSMIVPQKRLSASEKEKELCYVFEQWESDQVEVEHL 120
OY 121 ISQCHYOHGHINSYLPKMLQDRDTITLPAKGLDHIENLITSLDAKSLCAAEIVCKEY 180
Db 121 ISQCHYOHGHINSYLPKMLQDRDTITLPAKGLDHIENLITSLDAKSLCAAEIVCKEY 180
OY 181 RYISDGLMKKLLERWRTDSLMWGLAERGMGOYLFKNPPDGNAPNSFYALPKII 240
Db 181 RYISDGLMKKLLERWRTDSLMWGLAERGMGOYLFKNPPDGNAPNSFYALPKII 240
OY 241 ODITISNMRCGRHSIQRHCHSETSKGYCLODYDOKIVSGLRDNTIKIMDKNTLECK 300
Db 241 ODITISNMRCGRHSIQRHCHSETSKGYCLODYDOKIVSGLRDNTIKIMDKNTLECK 300
OY 301 RILGHTGSVLCLODYDERVYITGSSDSTVAVWVNTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILGHTGSVLCLODYDERVYITGSSDSTVAVWVNTGEMLNTLIHCEAVLHLRFNNGM 360
OY 361 VTCKSDRSIAVWMASTPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
Db 361 VTCKSDRSIAVWMASTPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
OY 421 EFVRTLNGHKGIGACLOYRDRLVYVSSSDNTIRLMDIECGACLVLEGHELVRCIRFDN 480
Db 421 EFVRTLNGHKGIGACLOYRDRLVYVSSSDNTIRLMDIECGACLVLEGHELVRCIRFDN 480
OY 481 KRIVSGAYDGKIKYWDLVAAIDPPAPAGTLCRLTVLHSGRVFLQDFEOLVSSSHDT 540
Db 481 KRIVSGAYDGKIKYWDLVAAIDPPAPAGTLCRLTVLHSGRVFLQDFEOLVSSSHDT 540
OY 541 ILIMDFLNDPAAQAEPSPSRITYTISR 569
Db 541 ILIMDFLNDPAAQAEPSPSRITYTISR 569

```

RESULT 12
AA878584
ID AAM78584 standard; Protein: 632 AA.

```

AC AAM78584;
XX 06-NOV-2001 (first entry)
XX Human protein SEQ ID NO 1246.
XX

```

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX WO200157190-A2.
 XX PD 09-AUG-2001.
 XX PF 05-FEB-2001; 2001WO-US04098.
 XX PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR N-PSDB; AAK51717.
 XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX PS Claim 20; Page 3505-3507; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAH80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAH80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX SQ Sequence 632 AA:
 Query Match 98.6%; Score 2992.5; DB 22; Length 632;
 Best Local Similarity 90.0%; Pred. No. 1.9e-282;
 Matches 569; Conservative 0; Mismatches 0; Indels 63; Gaps 1;
 QY 1 MDPAEAVIOEKALK-----14
 DB 1 MDPAEAVIOEKALKFMEFRSWCPGWNMTARSRLVATSTRVQCSMPSLWLGCSLADS 60
 QY 15 -----PANSSREDCCNGEPPRKTIIPKNSLRQTYNSCARLCLNQETV 57
 DB 61 MSLRLCLVPGTALTALEFNSSEREDCCNGEPPRKTIIPKNSLRQTYNSCARLCLNQETV 120
 QY 58 CLASTAMKTENCVAKTKLANGTSSMIVPKORLTASYEKEKLCVKEQMSQSESVQVFEV 117
 DB 121 CLASTAMKTENCVAKTKLANGTSSMIVPKORLTASYEKEKLCVKEQMSQSESVQVFEV 180
 QY 118 EHLISOMCHYGHINSYIKPMLQDFETALPARGLDIAENILSYLDAKSLCAAEVLCK 177
 DB 181 EHLISOMCHYGHINSYIKPMLQDFETALPARGLDIAENILSYLDAKSLCAAEVLCK 240

QY 178 EWTRYTSDGMLKKLIERVWRTDSLWRGLAERRGNGQYLFRKKPPDGNAPNSFYRALYP 237
 DB 241 EWTRYTSDGMLKKLIERVWRTDSLWRGLAERRGNGQYLFRKKPPDGNAPNSFYRALYP 300
 QY 238 KIIODIETIESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIVSGLDNTIKIMDKTL 297
 DB 301 KIIODIETIESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIVSGLDNTIKIMDKTL 360
 QY 298 ECKRILTGHTGSVLCQYDERVILITGSSDSYRVWDMVTGEMLNTLIHCEAVLHRENN 357
 DB 361 ECKRILTGHTGSVLCQYDERVILITGSSDSYRVWDMVTGEMLNTLIHCEAVLHRENN 420
 QY 358 GMAVTCCKDRSLAVNDMAFPDITLRVLYGHRANVNVDDDKYIVASGDRITKYWNT 417
 DB 421 GMAVTCCKDRSLAVNDMAFPDITLRVLYGHRANVNVDDDKYIVASGDRITKYWNT 480
 QY 418 STCEFEVRLNGHRKGIACQYDRSLVSGSSDNTIRLMDICGACLRVLEGEHEELVNCIR 477
 DB 481 STCEFEVRLNGHRKGIACQYDRSLVSGSSDNTIRLMDICGACLRVLEGEHEELVNCIR 540
 QY 478 FDNKRIVSGAYDGKIKVMDLVAAALDPAPAGTLCRLTVEHSGRVFRLQFDEQIVSSSH 537
 DB 541 FDNKRIVSGAYDGKIKVMDLVAAALDPAPAGTLCRLTVEHSGRVFRLQFDEQIVSSSH 600
 QY 538 DDTILIMFINDPAAQAPPPSPSKTYTISR 569
 DB 601 DDTILIMFINDPAAQAPPPSPSKTYTISR 632

RESULT 13

AAM79566
 ID AAM79566 standard; Protein: 654 AA.

AC AAM79566;
 XX

DT 06-NOV-2001 (first entry)
 XX

DE Human protein SEQ ID NO 3212.
 XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX WO200157190-A2.
 XX PD 09-AUG-2001.
 XX PF 05-FEB-2001; 2001WO-US04098.
 XX PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR N-PSDB; AAK52699.
 XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX

QY 118 EHLISOMCHYOHGINSYLAQMDPDTLPALPANGDHIENILSYDAKSLCAAEIYCK 177
 DB 203 EHLISOMCHYOHGINSYLAQMDPDTLPALPANGDHIENILSYDAKSLCAAEIYCK 262
 QY 178 EMYRTSDGMLMKKLIEMVRTDLSLWRGLAERGMGOYLFPKNPPDGNAPNSFYRALY 237
 DB 263 EMYRTSDGMLMKKLIEMVRTDLSLWRGLAERGMGOYLFPKNPPDGNAPNSFYRALY 322
 QY 238 KIIDDIETIESNMRCGRHSIQRHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTL 297
 DB 323 KIIDDIETIESNMRCGRHSIQRHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTL 382
 QY 298 ECKRLIHTGHTSVCLQYDERVITITSSDSSTVRYWVDNTGEMNTLHHCBAVLHLRFNN 357
 DB 383 ECKRLIHTGHTSVCLQYDERVITITSSDSSTVRYWVDNTGEMNTLHHCBAVLHLRFNN 442
 QY 358 GMMVTCSDRSIAVWMDASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 417
 DB 443 GMMVTCSDRSIAVWMDASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 502
 QY 418 STCEPVRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 477
 DB 503 STCEPVRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 562
 QY 478 FDNKRIYSGAYDGKIKVWDVLAALDPPAPAGTLCRLTIVHSGRVPRLODFEQIVSSSH 537
 DB 563 FDNKRIYSGAYDGKIKVWDVLAALDPPAPAGTLCRLTIVHSGRVPRLODFEQIVSSSH 622
 QY 538 DDTLLIMDFLNDPAQAEPSPSRITYTISR 569
 DB 623 DDTLLIMDFLNDPAQAEPSPSRITYTISR 654

RESULT 15
 ID AAM79568 standard; Protein; 654 AA.
 XX AAM79568;
 AC AAM79568;
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 3214.
 DE Human protein SEQ ID NO 3214.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 PN WO200157190-A2.
 PD 09-AUG-2001.
 PF 05-FEB-2001; 2001WO-US04098.
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 PA (HSE-) HYSEQ INC.
 PI Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 DR WPI; 2001-476283/51.

DR N-PSDB; AAK52701.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX Claim 20; Page 286-287; 6221pp; English.
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 CC Sequence 654 AA:
 SQ
 Query Match 98.5%; Score 2989.5; DR 22; Length 654;
 Best Local Similarity 89.9%; Pred. No. 3.9e-282;
 Matches 568; Conservative 1; Mismatches 0; Indels 63; Gaps 1;
 QY 1 MDPAAVLAQERALKFM----- 16
 DB 23 MDPAAVLAQERALKFMDFRSCPCQWNTMARSLRATSTSVGCSMPRSLMGCCSLADS 82
 QY 17 -----NSEREDCNNGEPKPKIPEKNSLQOTNSCARCLDNETV 57
 DB 83 MPELRCLYNPFGTGAITPAFQNSSEKEDCNNGEPKPKIPEKNSLQOTNSCARCLDNETV 142
 QY 58 CLASTAMKTENCVAATKLANGTSSMTIVPKORKLASYEKEKELCVKFFQWSESDQYEFV 117
 DB 143 CLASTAMKTENCVAATKLANGTSSMTIVPKORKLASYEKEKELCVKFFQWSESDQYEFV 202
 QY 118 EHLISOMCHYOHGINSYLAQMDPDTLPALPANGDHIENILSYDAKSLCAAEIYCK 177
 DB 203 EHLISOMCHYOHGINSYLAQMDPDTLPALPANGDHIENILSYDAKSLCAAEIYCK 262
 QY 178 EMYRTSDGMLMKKLIEMVRTDLSLWRGLAERGMGOYLFPKNPPDGNAPNSFYRALY 237
 DB 263 EMYRTSDGMLMKKLIEMVRTDLSLWRGLAERGMGOYLFPKNPPDGNAPNSFYRALY 322
 QY 238 KIIDDIETIESNMRCGRHSIQRHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTL 297
 DB 323 KIIDDIETIESNMRCGRHSIQRHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTL 382
 QY 298 ECKRLIHTGHTSVCLQYDERVITITSSDSSTVRYWVDNTGEMNTLHHCBAVLHLRFNN 357
 DB 383 ECKRLIHTGHTSVCLQYDERVITITSSDSSTVRYWVDNTGEMNTLHHCBAVLHLRFNN 442
 QY 358 GMMVTCSDRSIAVWMDASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 417
 DB 443 GMMVTCSDRSIAVWMDASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 502
 QY 418 STCEPVRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 477
 DB 503 STCEPVRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 562
 QY 478 FDNKRIYSGAYDGKIKVWDVLAALDPPAPAGTLCRLTIVHSGRVPRLODFEQIVSSSH 537
 DB 563 FDNKRIYSGAYDGKIKVWDVLAALDPPAPAGTLCRLTIVHSGRVPRLODFEQIVSSSH 622
 QY 538 DDTLLIMDFLNDPAQAEPSPSRITYTISR 569
 DB 623 DDTLLIMDFLNDPAQAEPSPSRITYTISR 654

Mon Feb 24 10:42:30 2003

us-09-601-168b-2.rag

Page 13

Search completed: February 20, 2003, 09:53:46
Job time : 89 secs

...

...

...

...

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:53:52 ; Search time 14 Seconds

(without alignments)
1685.716 Million cell updates/sec

Title: US-09-601-168b-2

Perfect score: 3034
Sequence: 1 MDPAAVLAQKALFKFMSSE.....PAAQAEPSPSPRTYTYISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3006	99.1	605	FW1A_HUMAN	Q9Y297 homo sapien
2	2597	85.6	518	TRCB_XENLA	Q91854 xenopus lae
3	2384.5	78.6	542	FW1B_HUMAN	Q9UBD1 homo sapien
4	1638.5	54.0	665	L123_CAEEL	Q09990 caenorhabdl
5	690	22.7	506	POF8_SCHPO	Q09855 schizosacch
6	590.5	19.5	605	POF1_SCHPO	P87053 schizosacch
7	575	19.0	678	SCOB_EMENTI	Q00659 emeticella
8	545	18.0	640	MT30_YEAST	P39014 saccharomyc
9	531.5	17.5	650	SCO2_NEUCR	Q01277 neosporea
10	520	17.1	579	SE10_CAEEL	Q93794 caenorhabdl
11	477.5	15.7	684	CC4_CANAL	P53699 candida alb
12	455.5	15.0	1356	HET1_PODAN	Q00808 podospora a
13	453	14.9	775	POF1_SCHPO	P87060 schizosacch
14	399	13.2	779	CC4_YEAST	P07834 saccharomyc
15	396.5	13.1	703	POP2_SCHPO	Q01277 neosporea
16	392	12.9	732	KMBB_DICDI	P90648 dictyostell
17	374	12.3	1526	YY46_ANASP	Q8Y111 anabaena sp
18	373.5	12.3	1258	YS00_ANASP	Q8Y112 anabaena sp
19	373.5	12.3	1683	YL24_ANASP	Q8Y113 anabaena sp
20	354	11.7	409	L1S1_HUMAN	P43034 homo sapien
21	354	11.7	409	L1S1_MOUSE	P43035 mus musculu
22	353	11.6	409	L1S1_BOVIN	P43033 bos taurus
23	341	11.2	515	YCM2_YEAST	P25382 saccharomyc
24	337	11.1	1146	KMHA_DICDI	P42527 dictyostell
25	334.5	11.0	422	FBW2_HUMAN	Q9UKT8 homo sapien
26	325.5	10.7	361	WDS_DROME	Q9V318 drosophila
27	324.5	10.7	334	WDR3_HUMAN	Q9UG99 homo sapien
28	318	10.5	422	FBW2_MOUSE	Q60584 mus musculu
29	313.5	10.3	376	YK14_CAEEL	Q17963 caenorhabdl
30	312.5	10.3	714	YUL2_YEAST	P47025 saccharomyc
31	307.5	10.1	742	PKMA_THRCU	P46955 thermomonos
32	307.5	10.1	1693	Y163_STYX3	Q5553 synechocyst
33	306	10.1	704	T2D4_DROME	P49846 drosophila

34	298.5	9.8	800	T2D4_HUMAN	Q1542 homo sapien
35	294.5	9.7	1249	APAF_RAT	Q9EPY5 rattus norv
36	292	9.6	659	YK16_YEAST	P36130 saccharomyc
37	290.5	9.6	614	TU11_SCHPO	Q09715 schizosacch
38	289	9.5	1249	APAF_MOUSE	Q88879 mus musculu
39	287.5	9.5	327	GBLP_BRANA	Q39336 brassica na
40	287	9.5	1248	APAF_HUMAN	Q14727 homo sapien
41	286.5	9.4	327	GBLP_ARATH	Q24456 arabidopsis
42	285.5	9.4	798	T2D4_YEAST	P38129 saccharomyc
43	283	9.3	473	PRP5_SCHPO	Q13615 schizosacch
44	283	9.3	713	TUPL1_YEAST	P16649 saccharomyc
45	281.5	9.3	682	TUPL1_KLDTA	P56094 kluyveromyc

ALIGNMENTS

RESULT 1	ID	FW1A_HUMAN	STANDARD:	PRT:	605 AA.
AC	Q9Y297	Q9Y213			
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-Trip)				
DE	(E3RSIXKAPPAB) (pikappabalpha-E3 receptor subunit).				
GN	FBXW1A OR FBW1A OR BTRCP OR BTRCP				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=99075339; PubMed=9859996;				
RA	Varon A., Hatzubai A., Davis M., Layon I., Amit S., Manning A.M.,				
RA	Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,				
RT	Identification of the receptor component of the Ikapabalpha-				
RT	ubiquitin ligase."				
RL	Nature 396:590-594(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	TISSUE=Lymphoid;				
RX	MEDLINE=98325370; PubMed=9660940;				
RA	Maroltin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,				
RA	Thomas D., Strebel K., Benarous R.;				
RT	A novel human WD protein, h-beta Trip, that interacts with HIV-1 Ypu				
RT	connects CD4 to the ER degradation pathway through an F-box motif."				
RL	Mol. Cell 1:565-574(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RX	MEDLINE=20003060; PubMed=10531035;				
RA	Cenciarrelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,				
RA	Pagano M.;				
RT	"Identification of a family of human F-box proteins."				
RL	Curr. Biol. 9:1177-1179(1999).				
RN	[4]				
RP	CHARACTERIZATION.				
RX	MEDLINE=99145464; PubMed=9990852;				
RA	Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,				
RA	Harper J.W.;				
RT	The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically				
RT	with phosphorylated destruction motifs in I-kappa-B-alpha and				
RT	beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro."				
RL	Genes Dev. 13:270-283(1999).				
CC	- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA				
CC	(PIKAPPABALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR				
CC	UBIQUITINATION AND DEGRADATION.				
CC	- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.				
CC	- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE				
CC	PRODUCED BY ALTERNATIVE SPLICING.				
CC	- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.				
CC	- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).				

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF101784; AAD08702.1; -
 DR EMBL: Y14153; CAA74572.1; -
 DR EMBL: AF129530; AAF04464.1; -
 DR Genew: HGNC:1144; BTRC.
 DR MIM: 603482; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR Prodom: PD000018; WD40; 4.
 DR SMART: SM00320; F-box; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS0181; F-box; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 6.
 DR PROSITE: PS00882; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 DR Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.
 KM DOMAIN 190 228 F-box.
 FT REPEAT 301 338 WD 1.
 FT REPEAT 341 378 WD 2.
 FT REPEAT 381 418 WD 3.
 FT REPEAT 424 461 WD 4.
 FT REPEAT 464 503 WD 5.
 FT REPEAT 505 541 WD 6.
 FT REPEAT 553 590 WD 7.
 FT REPEAT 590 605 WD 7.
 FT REPEAT 605 605 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 605 AA; 68866 MM; 4C67F3B7E400FD37 CRC64;

Query Match 99.1%; Score 3006; DB 1; Length 605;
 Best Local Similarity 94.0%; Pred. No. 7.7e-219;
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDPAAVLOKALKFM-----NSSREDC 24
 DB 1 MDPAAVLOKALKFM-----NSSREDC 60
 QY 25 NNGEPPKIIPEKNSLRQTYNSCARLCLNQEYCLASTAMKTENCYAKTKLANGTSSMIV 84
 DB 61 NNGEPPKIIPEKNSLRQTYNSCARLCLNQEYCLASTAMKTENCYAKTKLANGTSSMIV 120
 QY 85 PKORKLASAEKEKELCVKFEQWSSDOVEFEVHLISQCHYQGHINSYLYKPLQDF 144
 DB 121 PKORKLASAEKEKELCVKFEQWSSDOVEFEVHLISQCHYQGHINSYLYKPLQDF 180
 QY 145 ITPALPARGDHIENILSYDAKSLCAAEIYCKEYRYVSDGLMKKILERWRTDSLMR 204
 DB 181 ITPALPARGDHIENILSYDAKSLCAAEIYCKEYRYVSDGLMKKILERWRTDSLMR 240
 QY 205 GLAERGMGQYLKFNKPPDGNAPNSFYALYPKIIQDIETTESNRCGRHSIORIHCS 264
 DB 241 GLAERGMGQYLKFNKPPDGNAPNSFYALYPKIIQDIETTESNRCGRHSIORIHCS 300
 QY 265 ETSKGYVCLQYDQKIVSGIRDTIKIMDKNTLECKRIILTGHTGSVLCIYDERVITGS 324
 DB 301 ETSKGYVCLQYDQKIVSGIRDTIKIMDKNTLECKRIILTGHTGSVLCIYDERVITGS 360
 QY 325 SDSTVAVMDVNTGEMNLTLIHCEAVLHIFPNNGMVTCSKPSRLAVMDASPTDTTLNR 384
 DB 361 SDSTVAVMDVNTGEMNLTLIHCEAVLHIFPNNGMVTCSKPSRLAVMDASPTDTTLNR 420
 QY 385 VLVGHRAAVNVVDEDDKYIVSASGDTIKVWNTSTCEFTVINGKRGITACIYDRRLV 444
 DB 421 VLVGHRAAVNVVDEDDKYIVSASGDTIKVWNTSTCEFTVINGKRGITACIYDRRLV 480

QY 445 SGGSDNTILMIDIEGACALRVLEGHELVRCIRFDPNKRIVSGAYDGKIKVMDVLAALDPR 504
 DB 481 SGGSDNTILMIDIEGACALRVLEGHELVRCIRFDPNKRIVSGAYDGKIKVMDVLAALDPR 540
 QY 505 APAGTCLTLVHSGRVRLODFEPOIVSSSHDPTILMDPLNDPAAQAPRPSRSTY 564
 DB 541 APAGTCLTLVHSGRVRLODFEPOIVSSSHDPTILMDPLNDPAAQAPRPSRSTY 600

QY 565 TYISR 569
 DB 601 TYISR 605

RESULT 2
 TRCB_XENLA STANDARD; PRT; 518 AA.
 ID 091854; P70037; P70038;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-TrCP (Beta-transducin repeat-containing protein).
 GN FBXW1 OR BTRCP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93330289; PubMed=8393141;
 RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
 RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
 RT anaphase are rescued by Xenopus CDNs encoding N-ras or a protein
 RT with beta-transducin repeats."
 RL Mol. Cell. Biol. 13:4953-4966(1993).
 RN [2]
 RP SEQUENCE OF 302-518 FROM N.A.
 RX MEDLINE=97109804; PubMed=8952061;
 RA Hudson J.W., Alarcon V.B., Elinson R.P.;
 RT "Identification of new localized RNAs in the Xenopus oocyte by
 RT differential display PCR."
 RL Dev. Genet. 19:190-198(1996).
 CC - FUNCTION: Probably recognizes and binds to some phosphorylated
 CC proteins and promotes their ubiquitination and degradation.
 CC - SUBUNIT: PART OF A SCF (SKP1-CULIN-F-BOX) PROTEIN LIGASE COMPLEX
 CC (BY SIMILARITY).
 CC - DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
 CC MAURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
 CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
 CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
 CC TADPOLE EMBryo.
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC - SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC - SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M98268; AAA02810.1; -
 DR EMBL: U63922; AAB49671.1; -
 DR EMBL: U63922; AAB49672.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR Prodom: PD000018; WD40; 4.

DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 5.
 DR PROSITE; PS0082; WD_REPEATS_2; 7.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 DR Ubl conjugation pathway; Repeat; WD repeat.
 KW DOMAIN
 FT REPEAT 119 157 F-BOX.
 FT REPEAT 230 258 WD 1.
 FT REPEAT 270 298 WD 2.
 FT REPEAT 310 338 WD 3.
 FT REPEAT 353 381 WD 4.
 FT REPEAT 393 421 WD 5.
 FT REPEAT 433 461 WD 6.
 FT REPEAT 482 510 WD 7.
 FT REPEAT 510 518 WD 7.
 FT CONFLICT 302 304 GEM -> EFR (IN REF. 2).
 FT CONFLICT 516 518 GLA -> AAH (IN REF. 2).
 SQ SEQUENCE 518 AA: 59507 MW: 2A52EC19028127F3 CRC64;

Query Match 85.6%; Score 2597; DB 1; Length 518;
 Best Local Similarity 91.6%; Pred. No. 3, 9e-188;
 Matches 488; Conservative 7; Mismatches 8; Indels 30; Gaps 1;

QY 18 SSEREDCNNEPPKRIIPKRNLSKQYVNSCARLCLNDQVCLASTAMKTENCYAKTKLAN 77
 DB 13 ASEREDCNNEPPKRIIPKRIITKRLQ-----TKLAN 42
 QY 78 GTSMIVPKRKLSASKEKELCYKFEQMSDQVEFEVHLISOMCHYGHINSYK 137
 DB 43 GTSMIVPKRKLSANKEKELCYKFEQMSDQVEFEVHLISRCHYGHINSYK 102
 QY 138 PMLORDFITLPAAGLDHIAENILSYLDAKSLCAELVCKEYRVTSQGMKRLIERMV 197
 DB 103 PMLORDFITLPAAGLDHIAENILSYLDAKSLCAELVCKEYRVTSQGMKRLIERMV 162
 QY 198 RTDSLMGLAERKMGQYLRKRNPPDQNAOPNSFYRLYKIIQDITISNRCGRHSL 257
 DB 163 RTDSLMGLAERKMGQYLRKRNPPDQNAOPNSFYRLYKIIQDITISNRCGRHSL 222
 QY 258 ORHCRSETSKGVYCLQYDQOKIYSGLRDNTIKIMDNTECKRIITLGHGYSVLCLOYDE 317
 DB 223 ORHCRSETSKGVYCLQYDQOKIYSGLRDNTIKIMDNTECKRIITLGHGYSVLCLOYDE 282
 QY 318 RVITGSSDSIVRVWVDVNTGEMTLIHCEAVLHLRFNGMAMVTCSDRSIAVMDASP 377
 DB 283 RVITGSSDSIVRVWVDVNTGEMTLIHCEAVLHLRFNGMAMVTCSDRSIAVMDASA 342
 QY 378 TDITLRVYVGHRAAVVYVDFDDKYIYSASGDRITIKWNTSTCEFTVTLNGHKGICLQ 437
 DB 343 TDITLRVYVGHRAAVVYVDFDDKYIYSASGDRITIKWNTSTCEFTVTLNGHKGICLQ 402
 QY 438 YRDLRVYVSSSDNTIRLMDICGACLRVLEGEHELVCIRPDNKRIVSGAVDGKIKWDL 497
 DB 403 YRDLRVYVSSSDNTIRLMDICGACLRVLEGEHELVCIRPDNKRIVSGAVDGKIKWDL 462
 QY 498 VAALDPRAPACTICLIRLVHSGVFRLOQDEFQIVSSSHDDTILLMDFLNDP 550
 DB 463 VAALDPRAPACTICLIRLVHSGVFRLOQDEFQIVSSSHDDTILLMDFLNDP 515

RESULT 3
 FWLB_HUMAN STANDARD; PRT; 542 AA.
 AC Q9UKB1; Q9Y4C6; Q9P2S8; Q9P2S9;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-Trip2).
 GN FBXW1B OR FBXW1B OR BTBRC2 OR KIA00696.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID=9606;

RA [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20003060; PubMed=10531035;
 RA Cenciarelli C., Chiatur D.S., Guardavaccaro D., Parks W., Vidal M.,
 RA Pagano M.;
 RT "Identification of a family of human F-box proteins.";
 RL Curr. Biol. 9:1177-1179(1999).
 RN [2]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Fetal lung;
 RX MEDLINE=20160458; PubMed=10694485;
 RA Koike J., Sagera N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,
 RA Katoh M.;
 RT "Molecular cloning and genomic structure of the betaTRCP2 gene on
 RT chromosome 5q35.1.";
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98403880; PubMed=9734811;
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
 RA Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro."
 RL DNA Res. 5:169-176(1998).
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated
 CC proteins and promotes their ubiquitination and degradation.
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B AND C (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC
 CC EMBL; AF176022; AAF04528.1; -;
 CC EMBL; AB033279; BAA92329.1; -;
 CC EMBL; AB033280; BAA92330.1; -;
 CC EMBL; AB033281; BAA92331.1; -;
 CC EMBL; AB014596; BAA31671.1; ALT_INT.
 CC GeneW; HGNC:13607; FBXW1B.
 CC MIM; 605651; -;
 CC InterPro; IPR001810; F-box.
 CC InterPro; IPR001680; WD40.
 CC Pfam; PF00400; WD40; 7.
 CC Pfam; PF00646; F-box; 1.
 CC PRINTS; PR000320; GPROTEINBPT.
 CC ProDom; PD000018; WD40; 4.
 CC SMART; SM00256; FBOX; 1.
 CC SMART; SM00320; WD40; 7.
 CC PROSITE; PS50181; FBOX; 1.
 CC PROSITE; PS00678; WD_REPEATS_1; 5.
 CC PROSITE; PS0082; WD_REPEATS_2; 7.
 CC PROSITE; PS0294; WD_REPEATS_REGION; 1.
 CC Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.
 KW DOMAIN
 FT REPEAT 129 167 F-BOX.
 FT REPEAT 238 275 WD 1.
 FT REPEAT 278 315 WD 2.
 FT REPEAT 318 355 WD 3.
 FT REPEAT 361 398 WD 4.
 FT REPEAT 401 440 WD 5.
 FT REPEAT 442 478 WD 6.
 FT REPEAT 490 527 WD 7.
 FT VARSPIC 16 49 MISSING (IN ISOFORM A).
 FT VARSPIC 16 48 CSVPSLWLGACNLVSCALSCLOSMPSVRL -> NTSV

DB 203 NNNKRNKNTKTRINCSESKSGYCCQYDDDKIVSGLRDNTIKIMDKRKYSCSRILSGHT 262
 QY 308 GSVLCQYDERVITITGSSSTVAVMDVNTGEMLNTLIHCEAVLHFRFNNGMVATCSKDR 367
 DB 263 GSVLCQYDERVITITGSSSTVAVMDVNTGEMLNTLIHCEAVLHFRFNNGMVATCSKDR 322
 QY 368 STAVMDASPTDITLRLVGVHRAAVNVVDFDDKIVTASGDRITKVMSTCEVRYTLN 427
 DB 323 STAVMDASPTDITLRLVGVHRAAVNVVDFDDKIVTASGDRITKVMSTCEVRYTLN 382
 QY 428 GHRGACIQLQYRDLRVYSSSDMTIRLMDIEGACIARVIEGHELVRCIRFENKRIYVSGA 487
 DB 383 GHRGACIQLQYRDLRVYSSSDMTIRLMDIEGACIARVIEGHELVRCIRFENKRIYVSGA 442
 QY 488 YDKIKVMDVVALDPRAPAGTCLRTLVESGRVFLQDEFDEQVSSSHDDITILMDPL 547
 DB 443 YDKIKVMDVVALDPRAPAGTCLRTLVESGRVFLQDEFDEQVSSSHDDITILMDPL 502
 QY 548 NDPAQAEPSPSPRT 563
 DB 503 DAP-----PSGLPST 513

RESULT 5
 POFB_SCHPO STANDARD; PRT; 506 AA.
 ID 009855: 09P7V1:
 AC 01-FEB-1996 (Rel. 33, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE F-box/WD-repeat protein pof1.
 GN POF1 OR SPAC2966.01 OR SPAC30.05.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harrison C.L., Toda T.;
 RT "Systematic genome-wide analysis of F-box protein-encoding genes in
 RL fission yeast."
 RN Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RC SRRATN-972:
 RA MEDLINE-21848401; PubMed-11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
 RA Welteens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eber P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Kochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sipkavskii G., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe."

RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB061694; BAB55543.1; -
 DR EMBL: AL136538; CAB66464.1; -
 DR EMBL: Z66525; CA91423.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR ProDom: PD000018; WD40; 3.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 8.
 DR PROSITE: PS00181; PROX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 3.
 DR PROSITE: PS50082; WD_REPEATS_2; 6.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 70 116 F-BOX.
 FT REPEAT 219 256 WD 1.
 FT REPEAT 259 298 WD 2.
 FT REPEAT 301 338 WD 3.
 FT REPEAT 345 386 WD 4.
 FT REPEAT 388 426 WD 5.
 FT REPEAT 427 464 WD 6.
 FT REPEAT 468 505 WD 7.
 SQ SEQUENCE 506 AA; 58257 MW; CER34D4EFCBE2E10 CRC64;

Query Match 22.7%; Score 690; DB 1; Length 506;
 Best Local Similarity 30.4%; Pred. No. 1.4e-44;
 Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

QY 67 EMCVAK-----TLANGTSMIVPKOKISASYEKELCVKFFEQWSSDOVEPEHLIS 122
 DB 8 KNVSKVSDLTSCSDPSTSSPVPCLNPLS-----HNNRIDLRLDLA 50
 QY 123 OKCHYGHINSYKPMLODFTLAPAGLDIAENITSLYDAKSLCAELVCKEYRY 182
 DB 51 SLKSGVAVVYNNVSLTDFTEVFP-----EVSIRVFSYIDODLDICKKILMSKRMRL 106
 QY 183 TSDGMLMKRLI-----ERNVRDSDLMRG-----LAERGWG----- 213
 DB 107 LEDPGIKALYMQKGFVENVLEFEAWRTRHKFPQPFENFKLOQNLIGYGTMLPQ 166
 QY 214 QYLFKKPPDGNAPPSFRALYPKLIDILETIESMRCGRISLORINRS----- 264
 DB 167 QYLF-----DSNGRPLNMSYLY-----KEHALDSNWRGRPLVTFNPSIRPADQDF 217
 QY 265 -ETSKGVYCLQDDQKIVSGLRDNTIKIDKNTLLECKRLITGHTSVLCQYDER-VII 321
 DB 218 RALDSVYCEVQDDDEIMVSGSKDRIVSYVDVSRFLYLYLGHSSSVLCIDPCRRLNLY 277
 QY 322 TGGSSDSTVRVMDVNTGEMLNTLIHCEAVLHFRFNNGMVATCSKDRSIANV--DMASPTD 379
 DB 278 SGSSDSTIIMQWNRPLKVFEGHTDNLGVVSENYTSSSRHTAVWRDLATSPA 337
 QY 380 ITLRVLVGHRAAVNVVDDDK--YIVSASGDRITKVMSTCEVRYTLNGHKGRIACQ 437
 DB 338 ACN-HVLRGLHLSVNSVQSSKGLIVTASSDTTLTWITTGHCIRITHAQRKIAQ 396
 QY 438 YRDLVYSGSSDNTILMDIEGACIARVIEGHELVRCIRFENKRIYVSGAGDKIKYVMDL 497
 DB 397 YNCKFIYSGSSDITLIRIFASSGKILRLMIGHEDLIRTVRFNDKRIYSGCYGYRIRNN- 455

QY 498 VAALDPRAPAGTLCRTLVH-----SGRVFLQDFDERQIYSSHDDTLLWDF 546
 DB 456 -----FNTGECHCLVANSRNSKRVFLQDFHRIIACHTSSHTIWNF 497

RESULT 6
 ID POP1_SCHPO STANDARD; PRT; 605 AA.
 AC P87053;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE F-box/MD-repeat protein pop1 (SKP1-binding protein 1).
 GN POP1 OR SAK1 OR SPAC57A10.05c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Katayama S., Kitamura K., Toda T.;
 RT "Systematic genome-wide analysis of F-box protein-encoding genes in
 RT fission yeast."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Soutos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., O'Neill C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tilley A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
 RA Welbans I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Beyer P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Delaure V., Motlier S.,
 RA Galibert F., Aves S.D., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie M.R., Paulsen I., Potlakhin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe".
 RL Nature 415:871-880(2002).
 CC -I- FUNCTION: Probably recognizes and binds to some phosphorylated
 CC proteins and promotes their ubiquitination and degradation.
 CC -I- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX
 CC (BY SIMILARITY).
 CC -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -I- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AB032410; BAA84528.1; -

DR EMBL: Z94864; CAB08168.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PRODOM: PD000018; WD40; 5.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS00181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 2.
 DR PROSITE: PS00682; WD_REPEATS_2; 7.
 DR PROSITE: PS02944; WD_REPEATS_REGION; 1.
 KW Ubl conjugation; Repeat; WD repeat.
 FT DOMAIN 107 153
 FT REPEAT 271 299 WD 1.
 FT REPEAT 311 339 WD 2.
 FT REPEAT 350 379 WD 3.
 FT REPEAT 390 420 WD 4.
 FT REPEAT 432 460 WD 5.
 FT REPEAT 472 500 WD 6.
 FT REPEAT 510 538 WD 7.
 SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 19.5%; Score 590.5; DB 1; Length 605;
 Best Local Similarity 29.0%; Freq. No. 5.5e-37;
 Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

QY 80 SSMIVPKRKILASAYEKEKELCVKFEQWSE---SDQVEFVEHLISQCHVOHGINSYL 136
 DB 46 SSM-----HNEISGLESEKSGQREAVMAFSEASCSERIALOGILINSSSLSPASSTL 101
 QY 137 KMLQDRFTALPANGDLHIAENILSYDAKSLCAELVCKEWTYSDGMKRLIERM 196
 DB 102 DSLVRLDFSLTPV---EISFRILSFLDARSLSQAAQVSKRWELADDDVIMRMCEQH 157
 QY 197 VTTDSLMLGLAEFRGCMGYLFKN----- 219
 DB 158 INRK-----CEKCGKGLPLERNTLYAKASISQKRYERLTKRGVDAHSSPYKAKLD 211
 QY 220 -----KPDGNAPPNSE-----YRALPKTIIDIEITIESNMRCGRSLQ 258
 DB 212 DPTSSNETTISVYPPSPNSDSKFFLPFKTRPMKEVYAEKRC-----VECMWRGR 263
 QY 259 RHCH-----SESGVGYOLODDQKIVSGLDNTIKMDKTLDECKRLVGHSTVYCLQ 314
 DB 264 ---CQVYVLSGHSQVCLQVRLNLLASGYDATIRLMNLATFQVALLBEGHSGVYCLQ 320
 QY 315 YDERVITNGSSDSVTVRVMDVTGEMLNTLIHCEAVLHREFNMGMYTCKDRSIAVWM 374
 DB 321 EDCKLTGSGMDKTRIMNVRTSEGISLHGHDSVYCLTFPDSLLVSGSADCTVYKIMHF 380
 QY 375 ASPDITLRVLVGHRAAVNVDP--DDKYVVASGDRITKVNNTSCEVYRTLNGHKG 432
 DB 381 SGGRITLR-----GHTGPNVSVRIIRRGVLSSGSDSTIKIMSLFTNTCLHFSAHIGP 436
 QY 433 IACLQDRVLVSGSSDNTIRLMDIEGACGLRVLEGEHELRCIRPNKRIRVSAVYCKI 492
 DB 437 VQSLADLSRFLSCSLDTIKOMDIEKKKCVHTLFGHIEGWELADHLIRISAHGIV 496
 QY 493 KVMDLVAALDPRAPAGTLCRTLVHSGRVFLQDFEQIYSSHDDTLLIMFLNDP 550
 DB 497 KVMACE-----CVHTLKNHSBPVYSVALGCEVVGSGEDCKTYLMLFNAP 543

RESULT 7
 ID SCOB_EMENT STANDARD; PRT; 678 AA.
 AC Q00659;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-OCT-2001 (Rel. 40, Last annotation update)
 DE Sulfur metabolite repression control protein.

GN SCONB OR MAPBL.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 ON Eukaryota; Trichocomaceae; Emericella.
 RN NCBI_TaxId=5072;
 RP SEQUENCE FROM N.A.
 RA Natorff R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
 CC REPRESSION.
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U21220; AAC15905.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00464; F-box; 1.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR ProDom: PD000018; WD40; 4.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS00181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 4.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Transcription regulation; Repeat; WD repeat.
 FT DOMAIN 178 224 F-BOX.
 FT REPEAT 347 375 WD 1.
 FT REPEAT 387 415 WD 2.
 FT REPEAT 427 455 WD 3.
 FT REPEAT 466 496 WD 4.
 FT REPEAT 508 543 WD 5.
 FT REPEAT 553 595 WD 6.
 FT REPEAT 607 635 WD 7.
 FT REPEAT 647 675 WD 8.
 FT SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;
 Query Match 19.0%; Score 575; DB 1; Length 678;
 Best Local Similarity 26.9%; Pred. No. 9,4e-36;
 Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;

QY 354 RENNGMATYCKSKRSIAVWMASTPTDITLRVYVGHRAAVV--VDPDDKIYVSGSDRT 411
 DB 436 HFDASITLASGVYDKTVK IWNPEDKSTFLR-----GHTWMAVAVRVDTSRTVFSASDCT 491
 QY 412 IKVWNTSTCEFYRTLNH-----KRRIAC----- 435
 DB 492 VRLMDLTKTCIRTFHGVGOVYVPLPRFEFEHDEACENDLSTTSGDANPPIQA 551
 QY 436 -----LQYRD-----RLVYSGSDNTIRIMDEGCACRLVLEGHEL 472
 DB 552 SNGLEPNMAVSSQSAFSGSFNDGRAAPRYWTSALDSTIRIMETTRCRLTFPGHLEG 611
 QY 473 VQCIREFDNKRIYSGADGKIKYMDLVAALDPPAPATGLCLRTLVHSGRVRFLODFEQI 532
 DB 612 VVALGADTLRIYSGAEDRMKIMD-----PRTGCERTFTGHSPPVTCIGIDSRF 662
 QY 533 VSSSHDITLIMDF 546
 DB 663 ATGSEDCVRYMSF 676
 RESULT 8
 MT30_YEAST STANDARD: PRT: 640 AA.
 AC P39014;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE MET30 protein.
 GN MET30 OR YIL046W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 ON NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=42180-1A;
 RX MEDLINE=96069360; PubMed=8524217;
 RA Thomas D., Kuras L., Barbey R., Cherest H., Blaiseau P.L.,
 RA Surdin-Kerjan Y.;
 RT adenosylmethionine, is an essential protein with WD40 repeats*;
 RL Mol. Cell. Biol. 15:6526-6534(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
 CC GENES EXPRESSION.
 CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z46861; CA86905.1; -
 DR EMBL: L26505; AAA96717.1; -
 DR SGD: S0001308; MET30.
 DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR ProDom: PD000018; WD40; 4.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 6.
 DR PROSITE: PSS0181; FBOX; 1.
 DR PROSITE: PSS0678; WD_REPEATS_1; 4.
 DR PROSITE: PSS0082; WD_REPEATS_2; 6.
 DR PROSITE: PSS0294; WD_REPEATS_REGION; 1.
 DR Transcription regulation: Methionine biosynthesis;
 KW Cysteine biosynthesis: Repeat; WD repeat.
 FT DOMAIN 181 227 F-BOX.
 FT REPEAT 300 328 WD 1.
 FT REPEAT 340 368 WD 2.
 FT REPEAT 380 408 WD 3.
 FT REPEAT 419 449 WD 4.
 FT REPEAT 461 499 WD 5.
 FT REPEAT 509 538 WD 6.
 FT REPEAT 550 578 WD 7.
 FT REPEAT 607 635 WD 8 (POTENTIAL).
 FT CONFLICT 61 61 M -> I (IN REF. 1).
 SQ SOURCE 640 AA; 72835 MW; 5135D4BCA2E1B97 CRC64;

Query Match 18.0%; Score 545; DB 1; Length 640;
 Best Local Similarity 28.9%; Pred. No. 1.6e-33;
 Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

QY 34 IPEKNSLRQTYNSCARLCLN-QETVCLASTAMKTENCVAK-TKLANGSTSMIVPKQRKLS 91
 DB 86 LPEYNFTKFCFYRHNPDIQSPHTACYKQDLKRTQGINNNAIKLPLOESDIIHIIISKS 145
 QY 92 ASYEKKELCYKFEQMSSEDOVEFEHLISOMQHGHINSYKPMLODFETALPAR 151
 DB 146 NSMDKTRKL-----LIDGILSTISCFPLSYISLWYHMIKIPISILP- 188
 QY 152 GLDHIENLISYDAKSLCAELVCKEYRVTSQGLMKLIERVVRIDSLMRGLAER- 210
 DB 189 -QELSLKILSYLDCOSLONATRVCKKMKLADDDRVYHMEQOH-----DRKC 236
 QY 211 ---GMCQYLFKNK-----PPDGNAPNSFYALYPKIIQDIETTESNRCGRHSIQ 258
 DB 237 PNCGMGLPLHMKRARIQONSTGSSSNADIQOTTRPMKVIYREPRKVESNMKRG- 291
 QY 259 RIHCRESK-----GYCYQYDQKIVSGRLNTIKIMKNTLECKRIITGTSYLCIQ 314
 DB 292 --HCRIQEFGKHMVDVLTQFNTRILFTGSDYSTIGIMDLFTGKLIRLSGSHDGVKTL 349
 QY 315 YDERVITITSSDSSTVAVMDVNTGEMLNLHICFAVLHIFRNNGMVAVCSDKRSIAVDM 374
 DB 350 FDDRKLITGSLDKTRVWNYITGECISTYRGHSYLSVDSQKIVYSGADKIVKVMHV 409
 QY 375 ASPDTITLRVAVGHRAAVNVDFDKYI--VSASGDRTIKVMNTSGEYVFTLNG- 429
 DB 410 ESRCTYTLR-----GHEWVNCYKLPKFSFSCSDDTIRWMDIRINSCLKVFGHVGQ 465
 QY 430 KRGIAQLQYRD--RVVSSSS----- 448
 DB 466 VQKIIPILIKDENLNTDNTSDGSSPPDDPTMTGACADSDTPSPNOEFTVLNDENIPYTHL 525
 QY 449 -----DNTIRLMDIEGACGLAVLIEGHELVNCRIPDNKRIVSGAVDGKIKWDLAALDP 503
 DB 536 LSCGDNITIKLMDVTKICIRTFQGHVGVMDIADNFRIRISGSHDSIKWDLQSG- 582
 QY 504 RAPACTLIRLVHSGVFRIDPEFIYSSSDDTI 541
 DB 583 -----KCMHTF--NGR--RIQRETOHTQTOISLGDVY 609

AC 001277;
 DT 15-JUL-1998 (rel. 36, Created)
 DT 15-JUL-1998 (rel. 36, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Sulfur controller-2 (SCON2).
 GN SCON-2.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxId=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A;
 RX MEDLINE=95241499; PubMed=7724564;
 RA Kumar A, Paletta J V.;
 RT "The sulfur controller-2 negative regulatory gene of Neurospora
 crassa encodes a protein with beta-transducin repeats.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
 CC - FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
 CC - INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
 CC - SIMILARITY: CONTAINS 8 WD_REPEATS (TRP-ASP DOMAINS).
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC - SIMILARITY: BELONGS TO THE MET30/SCONE/SCON-2 FAMILY OF WD-REPEAT
 PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL: U17251; AAA68968.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR ProDom: PD000018; WD40; 4.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PSS0181; FBOX; 1.
 DR PROSITE: PSS0678; WD_REPEATS_1; 2.
 DR PROSITE: PSS0082; WD_REPEATS_2; 6.
 DR PROSITE: PSS0294; WD_REPEATS_REGION; 1.
 KW Transcription regulation: Repeat; WD repeat.
 FT DOMAIN 124 170 F-BOX.
 FT REPEAT 292 320 WD 1.
 FT REPEAT 332 360 WD 2.
 FT REPEAT 372 400 WD 3.
 FT REPEAT 411 441 WD 4.
 FT REPEAT 453 488 WD 5.
 FT REPEAT 528 564 WD 6.
 FT REPEAT 576 604 WD 7.
 FT REPEAT 616 644 WD 8.
 SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 17.5%; Score 531.5; DB 1; Length 650;
 Best Local Similarity 25.0%; Pred. No. 1.7e-32;
 Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;

QY 117 VEHLISOMCHYGHINSYKPMLODFETALPARGLDIAENILSYLDAKSLCAELVC 176
 DB 99 LOGILSQLEPQSLFVSREVENALKIDFISALPV-----ELAQVLCYLDYTSVLTAAQVS 154
 QY 177 KENVRYVTSQGLMKLIERVVRIDSLMRGLAERGGQYLFKNKP-----P 222
 DB 155 QKRPITLADSDAVVWVWVWCEQHVNRK-----CTKCGWGLPLERKILNRYTRQRLAKGP 208
 QY 223 DGNA----- 234
 DB 209 QGRVTELADSHDSQDSVNHQGRPAALAEEDDPKKRQCMMAAASKAIVTOPKTRSWKA 268

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
 CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
 CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
 CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
 CC VARIOUS ASPECTS OF SPOGULATION. REQUIRED FOR HTAI-HTBI LOCUS
 CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X96763; CAA65538.1; -
 DR InterPro: IPR001810; F-box.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PRODOM: PD000018; WD40; 1.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 4.
 DR PROSITE: PS50082; WD_REPEATS_2; 4.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
 FT DOMAIN 212 258
 FT REPEAT 322 351 WD 1.
 FT REPEAT 363 391 WD 2.
 FT REPEAT 403 431 WD 3.
 FT REPEAT 442 468 WD 4 (POTENTIAL).
 FT REPEAT 478 506 WD 5.
 FT REPEAT 519 549 WD 6.
 FT REPEAT 561 589 WD 7.
 SO SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;
 Query Match 15.7%; Score 477.5; DB 1; Length 684;
 Best Local Similarity 25.4%; Pred. No. 2.1e-28;
 Matches 136; Conservative 95; Mismatches 184; Indels 121; Gaps 16;

QY 446 GSSDNTIRLMIEGACGLVLEGHELYNCIRFDNRKRYSGAYDGKIRKWDVLAALDPA 505
 DB 579 AADATLRITMDAKTGELRSKLGHGALTCFPHGDGLRVVSGS - EKMLTMV----- 629
 QY 506 PAGTLCRLTVSHSGRVFLQD-----EQIYSSHDITIL-TWPLNDPAQ 553
 DB 630 EKGTRARLLSDVTGIGWQVDFYRKYAAVGRILNEDEGETFIELDF-SQPLQ 684
 RESULT 12
 ID HET1_PODAN STANDARD. PRT: 1356 AA.
 AC 000808;
 DT 15-JUL-1998 (rel. 36, Created)
 DT 15-JUL-1998 (rel. 36, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Vegetable Incompatibility protein HET-E-1.
 GN HET-E1.
 OS Podospora anserina.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Lasiosphaeriaceae; Podospora.
 OX NCBI_TaxID=5145;
 RN [1]
 RP MEDLINE=96009891; PubMed=7557402;
 RA Saude S., Turc B., Bequeret J.;
 RT "A gene responsible for vegetative incompatibility in the fungus
 RT Podospora anserina encodes a protein with a GTP-binding motif and G
 RT beta homologous domain."
 RL Gene 162:135-139(1995).
 CC -1- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
 CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
 CC HET-C.
 CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: CONTAINS 1 NACHT DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L28125; AAA85775.1; -
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 10.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PRODOM: PD000018; WD40; 10.
 DR SMART: SM00320; WD40; 10.
 DR PROSITE: PS50837; NACHT; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 10.
 DR PROSITE: PS50082; WD_REPEATS_2; 10.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW GTP-binding; Repeat; WD repeat.
 FT DOMAIN 294 629
 FT NP_BIND 300 307 GTP (POTENTIAL).
 FT REPEAT 839 869 WD 1.
 FT REPEAT 881 911 WD 2.
 FT REPEAT 923 953 WD 3.
 FT REPEAT 965 995 WD 4.
 FT REPEAT 1007 1037 WD 5.
 FT REPEAT 1049 1079 WD 6.
 FT REPEAT 1091 1121 WD 7.
 FT REPEAT 1133 1163 WD 8.
 FT REPEAT 1175 1205 WD 9.
 FT REPEAT 1217 1247 WD 10.
 SO SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;
 Query Match 15.0%; Score 455.5; DB 1; Length 1356;
 Best Local Similarity 26.0%; Pred. No. 2.3e-26;
 Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

QY 28 EPPRKIIPEKNSLRQTYNSCAR-ICLNQETVCLASTAMK-----TENCYAKTKLANGTSS 81
 DB 581 EMDVDSDDESELEIYKLCGSLIIRERIVYVHOSAKPFLLTGTSADKSNKASQSEAF 640
 QY 82 MIVPR-----ORIKSASYEKEKE--LCYKPEBMSQDQVEVEHL--ISQMCHY 127
 DB 641 LVEPTGIEDVSYILFWRSILNVMQSKLRDRDYCLINAPFLIDNRPDPPLATVRYSCIY 700
 QY 128 OHGHINYLKP-----MLQRO-----FT-----ALPANGLDHI--AEN 159
 DB 701 WIDHLRDLVSTSSKWHLLQDDGDHHRFLTYKLYLWLEALSLRLAP--EGVIAHROLES 759
 QY 160 ILASYLDKSCIAAEIYCKEYRVYSDGMLMKLIER---NVRTDSLMRGIAERRGMOYL 216
 DB 760 LIGTINGRLIA---IYRDCYRF---ALSTRMIEKAPLQAYTSAL-----V 800
 QY 217 FKKNPPDGNAPNPSFYALY---PKLIIDILETIESWRCGRHSL----- 257
 DB 801 F-----APTDSMIKKIKKEPQWISITISVEAEEMNACTOTLEGHGSVLSVAFSAD 852
 QY 258 -QRIHCS-----ETSKG-----VYCLQY--DDQKIVSGLRDNTIKIMDK 294
 DB 853 GQRYASGSDDKTIKIMDTASGCTQTLLEGHGSVSAFSPDRERVASGSDDKTIKWDA 912
 QY 295 NTECKRILGHTGVSCLQY--DERVLIIGSSDSIVVMDVNTGEMLNTLIHHCFAVLH 352
 DB 913 ASGCTQTLLEGHGRVQSVAFSPDQGRVASSGSDHTIKIMDASGCTQTLLEGHGSVLS 972
 QY 353 LRFN-NGMWTYC-SKDSIAVMDASPPDITLRVLYVGHRAVNVVF--DDKIYVASG 408
 DB 973 VAFSPDQGRVASSGSDHTIKIMDTASG--TCQTLLEGHGSVSAFSPDQGRVASSGSD 1029
 QY 409 DRTIKVWNTSTCEVPTLNGHKRGIACTQYR--DRLVVGSSSDTIFLMDIEGCACRVL 466
 DB 1030 DKTIKIMDTASGCTQTLLEGHGSVQSVSPDQGRVASSGSDHTIKIMDASGCTQTL 1089
 QY 467 EGHLELVNCFIR--DNKRIYSGAYDGKIKWDLVLAALDPRAACTICLRLIVHSGNFR 524
 DB 1090 EGHGDSVMSVAFSPDQGRVASSGSDHTIKIMD-----AASGT-CTQTLLEGHGVHS 1140
 QY 525 LQF--DERQIVSSSHDTITLIMDLNDPAAQ 553
 DB 1141 VAFSPDQGRVASSGSDHTIKIMDASGCTQ 1171
 RESULT 13
 POP1_SCHPO STANDARD; PRT; 775 AA.
 AC P87060;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WD-repeat protein popl.
 GN POP1 OR SPBC262.18.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 OC NCHI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA MEDLINE=97347242; PubMed=9203581;
 RA Komhiani K., Toda T.;
 RT "Fission yeast WD-repeat protein popl regulates genome ploidy through
 RT ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1
 RT and the S-phase initiator Cdc18.";
 RL Genes Dev. 11:1548-1560(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Beal N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jags K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolkeart G., Aert R., Koben J., Grynopre B.,
 RA Weljens I., Vansteels E., Rieger M., Schefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Beyer P., Zimmermann W., Wedler H., Mambutt R., Punelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerniti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 RN [3]
 RP SUBUNIT.
 RC STRAIN-972;
 RX MEDLINE=99144318; PubMed=9990507;
 RA Komhiani K.-I., Ochotoreia I., Toda T.;
 RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
 RT complexes together with cullin-1 in fission yeast SCF (Skip-cullin-1-
 RT F-box) ubiquitin ligase.";
 RL Genes Cells 3:721-735(1998).
 CC - FUNCTION: Involved in maintenance of ploidy through proteasome
 CC dependent degradation of CDK inhibitor rum1 and S-phase initiator
 CC cdc18. Functions as a recognition factor for rum1 and cdc18, which
 CC are subsequently ubiquitinated and targeted to the 26S proteasome
 CC for degradation.
 CC - SUBUNIT: Homodimer and heterodimer with pop2. Binds to pop1 and
 CC cdc18.
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC - SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 DR EMBL: Y08391; CA65671.1;
 DR EMBL: AL022103; CA17898.1;
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR Pfam: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR ProDom: PD000018; WD40; 3.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 6.
 DR PROSITE: PS00181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 3.
 DR PROSITE: PS50082; WD_REPEATS_2; 4.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat. WD repeat.
 FT DOMAIN 298 345 F-BOX.
 FT REPEAT 444 472 WD 1.
 FT REPEAT 484 538 WD 2.
 FT REPEAT 575 603 WD 3.
 FT REPEAT 615 645 WD 4.

```

RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972:
RX MEDLINE=95400292; PubMed=7670463;
RA Murkanti Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasunuma S.-I., Sasunuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT *Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae".
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 1-579 FROM N.A.
RC STRAIN=S288c / AB972:
RA Barrell B.G., Churcher C., Rajandream M.A.;
RL Submitted (SEP-1994) to the EMBL/Genbank/DBJP databases.
CC -! FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONFIGURATION AND
CC VARIOUS ASPECTS OF SPOKULATION. REQUIRED FOR HTAI-HTB1 LOCUS
CC TRANSCRIPTION ACTIVATION.
CC -! SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -! SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05625; CAA29113.1; -
DR EMBL; D50617; BAA09229.1; -
DR EMBL; Z46255; CAA86341.1; -
DR PIR; A26867; A26867.
DR SGD; S000185; CDC4.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001880; WD40.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 4.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT FT DOMAIN 272 319 F-BOX.
FT REPEAT 380 408
FT REPEAT 420 449 WD 1.
FT REPEAT 461 493 WD 2.
FT REPEAT 528 556 WD 3.
FT REPEAT 568 598 WD 4.
FT REPEAT 630 658 WD 5.
FT REPEAT 669 698 WD 6.
FT REPEAT 722 741 WD 7.
FT CONFLICT 460 460 K -> E (IN REF. 1).
SQ SEQUENCE 779 AA; 86089 MW; 0348F2F8FA78F3BC CRC64;

Query Match 13.2%; Score 399; DB 1; Length 779;
Best Local Similarity 24.3%; Pred. No. 2e-22;
Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCLASTAMKTENCVAKTKLANGTSSMIVPKRKLASYSERK-----ELCKYFEQW 108
DB 194 TTPLAKTKTKINN-----NNNTADLIESKDSITISEYLSDEFSAINNNTLPHAYRK-- 244

QY 109 SESDQVEFEVHLSQMHCHYGHINSVLKPLQLODFLTALPARLDHIAENILSYDAKS 168
DB 245 -----NLIRLVANMRSELSDLGITLKNLKHDLTSLP-----ETSLKIFNYLODF 294

Y 169 LCAELVCKEWMYR-VTSDGMLMKRLI--ERRVVRDLSLRLGLAERGGQYLFRNKKPPDGN 225

```

DB 295 IINSLGVSNMNNKIRKSTSLMKLISENFV-----SPKGF 331
 QY 226 APPNSFYRALYPIID-----IEFT-IESNMGCRSLORINCRSETSKGVCLQYD 276
 DB 332 NSINLKISQKYPKLSOODRLRLSFLSENIIFLKWMYNPKFPPQPTLLRGHMTSVITCLOPE 391
 QY 277 DQKIVSGLRNRTIKINDKNTLECKRLLTGHGTVCLQYD-RVITGSSDSIVRWADYN 335
 DB 392 DNVTITGADGDKMLRYVDISINKKFLQLSGHGVWMLKTAAGGILVSGSDTRVIRWADIK 451
 QY 336 TGEMLNTLIHHCBAVLHLRFNNGMATCSKRSIAVWDASPDITLRRVLVGHRAVAV 395
 DB 452 KG-----CT-----HFVKGHNSTVAC 468
 QY 396 VEDD-----KYIVSASGDRITKWMNT-----STCE-----FVPTLNG 428
 DB 469 LDIVETKNIKVIYTGSRDNTLHWKLPKSSVDPDGEHEDYPLVFTPTPEPNPFVGVCLNG 528
 QY 429 HKGICACLOVRDLVYVSGSDNTIRLMDIEGACLVLEGHEELVNCIRFDN--KRIYVG 486
 DB 529 HMASVRIYSGHGNIVYSGSYDNTLIWDAOMKCLYLISGHDRIYSTYDHERKRCISA 588
 QY 487 AYGKIKVMDL-----VAALDPRAPAGTL--CLRTLYEHSCHVFLQDEFEQIVSS 535
 DB 589 SMDTTRIMDLNEMNNGECSTVITNSAPCAKILGMYTLQGHVALGLRLSDKFLVSA 648
 QY 536 SHDDTLIMDLNDPAAOAEPSPSPRYTY 566
 DB 649 AADGSRGWD-AND-----YSRKFSY 668

RESULT 15

POP2_SCHPO STANDARD: PRT: 703 AA.

AC 014170;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WD-repeat protein Pop2 (Proteolysis factor sud1).
 GN POP2 OR SUD1 OR SPAC4D7.03.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=99227353; PubMed=10209119;
 RA Wolf D.A., McKeon F., Jackson P.K.;
 RT "F-box/WD-repeat proteins Pop1 and Sud1/Pop2 form complexes that
 RT bind and direct the proteolysis of cdc18p.";
 RL Curr. Biol. 9:373-376(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98318628; PubMed=9653157;
 RA Jallepalli P.V., Tien D., Kelly T.J.;
 RT "Sud1 targets cyclin-dependent kinase-phosphorylated Cdc18 and Rnn1
 RT proteins for degradation and stops unwanted diploidization in fission
 RT yeast.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21648401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Mblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
 RA Woltjens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hiltbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mortier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Dague R.R., Cruzado L., Ulmeyer J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovsky G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 RN [4]
 RP FUNCTION, AND SUBUNIT.

RC STRAIN=972;
 RX MEDLINE=99144318; PubMed=9990507;
 RA Komlani K.-I., Ochotorena I., Toda T.;
 RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
 RT complexes together with cullin-1 in fission yeast SCF (Skip-cullin-1-
 RT F-box) ubiquitin ligase.";
 RL Genes Cells 3:721-735(1998).
 CC -1- FUNCTION: Involved in maintenance of ploidy through proteasome
 CC dependent degradation of CDK inhibitor Rnn1 and S-phase initiator
 CC cdc18. Functions as a recognition factor for Rnn1 and cdc18, which
 CC are subsequently ubiquitinated and targeted to the 26S proteasome
 CC for degradation.
 CC -1- SUBUNIT: Homodimer and heterodimer with pop1. Binds to pcu1 and
 CC cdc18.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF038867; AAC95480.1; -
 CC EMBL: AF064515; AAC99496.1; -
 CC EMBL: Z98602; CAB11275.1; -
 CC InterPro: IPR001810; F-box.
 CC InterPro: IPR001680; WD40.
 CC Pfam: PF00400; WD40; 7.
 CC Pfam: PF00646; F-box; 1.
 CC PRINTS: PR00320; GPROTEINRPT.
 CC PRODOM: PD000018; WD40; 3.
 CC SMART: SM00256; FBOX; 1.
 CC SMART: SM00320; WD40; 6.
 CC PROSITE: PS00181; FBOX; 1.
 CC PROSITE: PS00678; WD_REPEATS_1; 3.
 CC PROSITE: PS00682; WD_REPEATS_2; 3.
 CC PROSITE: PS00294; WD_REPEATS_REGION; 1.
 CC KW Repeat; WD repeat.
 CC FT DOMAIN 236 283 F-BOX.
 CC FT REPEAT 389 417 WD 1.
 CC FT REPEAT 429 473 WD 2.
 CC FT REPEAT 505 533 WD 3.
 CC FT REPEAT 545 575 WD 4.
 CC FT REPEAT 587 615 WD 5.
 CC FT REPEAT 625 654 WD 6.
 CC SEQUENCE 703 AA; 79605 MW; 0336A0568C152396 CRC64;

Query Match 13.1%; Score 396.5; DB 1; Length 703;
 Best Local Similarity 23.3%; Pred. No. 2.7e-22;

[illegible]

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 22:09:15 ; Search time 99 Seconds
(without alignments)
6663.250 Million cell updates/sec

Title: US-09-601-168b-1

Perfect score: 2151
Sequence: 1 tgcgttgctgcgcgcctgac.....gtttgcagagaaaaaaa 2151

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/BCRUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124.4	5.8	126	US-09-172-841-2	Sequence 2, App1
2	82.2	3.8	2237	US-08-914-999-7	Sequence 7, App1
3	67.2	3.1	2481	US-08-899-578-1	Sequence 1, App1
4	50.8	2.4	2085	US-08-283-917-8	Sequence 8, App1
5	50.8	2.4	2085	US-08-961-716-8	Sequence 8, App1
6	50	2.3	7218	US-08-232-463-14	Sequence 14, App1
7	47.8	2.2	2152	US-08-188-582-17	Sequence 17, App1
8	47.8	2.2	2152	US-08-646-715-17	Sequence 17, App1
9	39.8	1.9	7218	US-08-232-463-14	Sequence 14, App1
10	38.6	1.8	433	US-08-664-596B-1	Sequence 1, App1
11	38.6	1.8	433	US-08-738-367-1	Sequence 1, App1
12	38.6	1.8	733	US-08-738-367-6	Sequence 6, App1
13	38.6	1.8	3380	US-09-156-425-1	Sequence 1, App1
14	36.6	1.7	1613	US-08-965-048-3	Sequence 3, App1
15	36.6	1.7	18596	US-09-318-448-11	Sequence 11, App1
16	36.6	1.7	45989	US-08-965-048-6	Sequence 6, App1
17	36.2	1.7	7886	US-08-751-189-2	Sequence 2, App1
18	36.2	1.7	7886	US-09-060-836-2	Sequence 2, App1
19	36.2	1.7	7886	US-09-184-445-2	Sequence 2, App1
20	35.8	1.7	3434	US-09-605-785-476	Sequence 476, App
21	35.8	1.7	3434	US-09-439-313-476	Sequence 476, App
22	35.8	1.7	5829	US-09-605-785-473	Sequence 473, App
23	35.8	1.7	5829	US-09-439-313-473	Sequence 473, App
24	35.4	1.6	7881	US-08-751-189-1	Sequence 1, App1
25	35.4	1.6	7881	US-09-060-836-1	Sequence 1, App1
26	35.4	1.6	7881	US-09-184-445-1	Sequence 1, App1
27	35.4	1.6	10881	US-09-357-206A-9	Sequence 9, App1

28	35.4	1.6	12286	4	US-09-357-206A-1	Sequence 1, App1
29	34.8	1.6	3465	4	US-08-914-999-5	Sequence 5, App1
30	34.6	1.6	1611	4	US-09-302-769-13	Sequence 13, App1
31	34.6	1.6	1848	1	US-08-313-553-10	Sequence 10, App1
32	34.6	1.6	1848	3	US-08-767-993-10	Sequence 10, App1
33	34.6	1.6	2246	4	US-09-032-742-3	Sequence 3, App1
34	34.6	1.6	2246	4	US-09-032-742-21	Sequence 21, App1
35	34.6	1.6	2246	4	US-09-032-742-22	Sequence 22, App1
36	34.6	1.6	2246	4	US-09-032-742-24	Sequence 24, App1
37	34.6	1.6	2246	4	US-09-032-742-25	Sequence 25, App1
38	34.6	1.6	5558	4	US-08-961-527-103	Sequence 103, App
39	34.4	1.6	649	4	US-09-385-982-336	Sequence 336, App
40	34.2	1.6	3747	4	US-09-690-364-17	Sequence 17, App1
41	34.2	1.6	7042	4	US-09-092-508-1	Sequence 1, App1
42	34.2	1.6	7042	4	US-09-435-115-1	Sequence 1, App1
43	34.2	1.6	7042	4	US-09-098-310-1	Sequence 1, App1
44	34.2	1.6	7042	4	US-09-690-364-21	Sequence 21, App1
45	34.2	1.6	7075	4	US-09-092-508-15	Sequence 15, App1

ALIGNMENTS

```

RESULT 1
US-09-172-841-2
; Sequence 2, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; APPLICANT: Elledge, Stephen J.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT APPLICATION NUMBER: US/09/172,841
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 126
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-172-841-2

Query Match          5.8%; Score 124.4; DB 4; Length 126;
Best Local Similarity 99.2%; Pred. No. 6.2e-31;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 511 CTGCCAGCTCGGGGATGGATCATATCGCTGAGAACATTTCTCATPACCTGGATGCCAAA 570
    |||||||
DB 1 CTGCCAGCTCGGGGATGGATCATATTCCTGAGAACATTTCTCATPACCTGGATGCCAAA 60

QY 571 TCACATGTCGTCGCTGACATTTGTCGACGAAGATGTCACGAGTACCGTTCATGGATG 630
    |||||||
DB 61 TCACATGTCGTCGCTGACATTTGTCGACGAAGATGTCACGAGTACCGTTCATGGATG 120

QY 631 CTGTGG 636
    |||||
DB 121 CTGTGG 126

RESULT 2
US-08-914-999-7
; Sequence 7, Application US/08914999
; Patent No. 6346406
; GENERAL INFORMATION:
; APPLICANT: Ryazanov, Alexey G.
; APPLICANT: Hailt, William N.
; APPLICANT: Pavut, Karen S.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; NUMBER OF SEQUENCES: 25
; METHODS OF USE THEREFOR
; CORRESPONDENCE ADDRESS:

```


Query Match	2.4%	Score 50.8;	DB 2;	Length 2085;
Best Local Similarity	47.3%;	Pred. No. 5.6e-06;		
Matches 262;	Conservative 0;	Mismatches 277;	Indels 15;	Gaps 3

```

1 TITLE OF INVENTION: RECOMBINANT FOMLPOX VIRUS
2
3 NUMBER OF SEQUENCES: 52
4
5 CORRESPONDENCE ADDRESS:
6
7 ADDRESSEE: Foley & Lardner
8
9 STREET: 1800 Diagonal Road, Suite 500
10
11 CITY: Alexandria
12
13 STATE: VA
14
15 COUNTRY: USA
16
17 ZIP: 22313-0299
18
19 COMPUTER READABLE FORM:
20
21 MEDIUM TYPE: Floppy disk
22
23 COMPUTER: IBM PC compatible
24
25 OPERATING SYSTEM: PC-DOS/MS-DOS
26
27 SOFTWARE: Patent In Release #1.0, Version #1.25
28
29 CURRENT APPLICATION DATA:
30
31 APPLICATION NUMBER: US/08/232,463
32
33 FILING DATE:
34
35 CLASSIFICATION: 435
36
37 PRIOR APPLICATION DATA:
38
39 APPLICATION NUMBER: US/07/935,313
40
41 FILING DATE:
42
43 APPLICATION NUMBER: EP 91 114 300.6
44
45 FILING DATE: 26-AUG-1991
46
47 ATTORNEY/AGENT INFORMATION:
48
49 NAME: BENT, Stephen A.
50
51 REGISTRATION NUMBER: 29,768
52
53 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
54
55 TELECOMMUNICATION INFORMATION:
56
57 TELEPHONE: (703)836-9300
58
59 TELEFAX: (703)683-4109
60
61 TELEX: 899149
62
63 INFORMATION FOR SEQ ID NO: 14:
64
65 SEQUENCE CHARACTERISTICS:
66
67 LENGTH: 7218 base pairs
68
69 TYPE: nucleic acid
70
71 STRANDEDNESS: single
72
73 TOPOLOGY: linear
74
75 IMMEDIATE SOURCE:
76
77 CLONE: pTZgpt-Fls
78
79 US-08-232-463-14
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:55:06 ; Search time 22 seconds

(without alignments)
2486.386 Million cell updates/sec

Title: US-09-601-168b-2

Perfect score: 3034

Sequence: 1 MDPAEAVLQERKALRFMNSSE.....PAAQAEPPSPSRITYTISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR-73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2597	85.6	518	2	B48088
2	1635.5	53.9	701	2	T16607
3	690	22.7	506	2	T50211
4	590.5	19.5	605	2	T38932
5	545	18.0	640	2	S49932
6	531.5	17.5	650	2	T46660
7	520	17.1	579	2	T22703
8	519.5	17.1	267	2	S62507
9	455.5	15.0	1356	2	T18521
10	453	14.9	775	2	T45136
11	413.5	13.6	1227	2	AE1810
12	399	13.2	779	2	S56245
13	396.5	13.1	703	2	T43557
14	378.5	12.5	1189	2	A12493
15	375	12.4	1526	2	AC1842
16	374	12.3	1526	2	AC2329
17	373.5	12.3	1258	2	A12155
18	373.5	12.3	1683	2	AE2071
19	361.5	11.9	677	2	AE1861
20	358.5	11.8	559	2	AB2202
21	354	11.7	409	2	S61113
22	354	11.7	410	2	S48052
23	341	11.2	515	2	S19487
24	339.5	11.2	777	2	T41075
25	337	11.1	1146	2	A55532
26	336	11.1	676	2	AH2195
27	334	11.0	589	2	AG2400
28	333.5	11.0	1711	2	AD1842
29	332.5	11.0	317	2	T46032

30	330	10.9	333	2	G85034	probable WD-repeat
31	320	10.5	934	2	AG1889	WD-40 repeat prote
32	318	10.5	422	2	A56640	CDC4 repeat unit-c
33	313.5	10.3	376	2	T19266	hypothetical prote
34	312.5	10.3	714	2	S56893	hypothetical prote
35	312	10.3	786	2	AC2375	WD-40 repeat-prote
36	312	10.3	876	2	T51507	WD-40 repeat protei
37	307.5	10.1	1693	2	S76086	beta transducin-11
38	306.5	10.1	598	2	AE2415	WD-repeat protein
39	306	10.1	1189	2	AH2154	WD-repeat protein
40	305	10.1	704	2	S33263	transcription init
41	299.5	9.9	1194	2	T03818	apoptotic proteina
42	298	9.8	265	2	AF1890	WD-repeat protein
43	297.5	9.8	357	2	A12099	WD-40 repeat prote
44	297	9.8	502	2	T41148	trip-asp repeat con
45	296.5	9.8	304	2	AG1837	WD-40 repeat prote

ALIGNMENTS

RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-trcp

C:Species: Xenopus laevis (African clawed frog)

C>Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000

C:Accession: B48088

R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: B48088; MUID:9330289; PMID:8393141

A:Accession: B48088

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SPe>

A:Cross-references: GB:M98268; NID:9295542; PIDN:AAA02810.1; PID:9295543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <WD1>

Query Match	Score	Length	DB ID	Description
Best Local Similarity	91.6%	Pred. No. 1.2e-193		
Matches 488; Conservative 7; Mismatches 8; Indels 30; Gaps 1;				
18	SSEREDCNNGEPKRIIPEKNSLRQTYNSCARLQNETVCLASTAMKTENCVAKTLAN	77		
13	ASEREDCNNDPEPKRIITEKTLRQ-----TKLAN	42		
78	GTSMIVPKRKRLISASYEKEKELCVKFFQMSRSDQVEVEHLISOMCHYOHGHTSYLK	137		
43	GTSMIVPKRKRLISANKEKEKELCVKFFQMSRSDQVEVEHLISOMCHYOHGHTSYLK	102		
138	PMIORDFTALPARGDHIENITLSYDAKSLCAELVCKENRYATSDGLMKLTERVY	197		
103	PMIORDFTALPARGDHIENITLSYDAKSLCAELVCKENRYATSDGLMKLTERVY	162		
198	RTDSLWGLAERKRGWGYLFKNRPDGNAPNSFYRLPKIITQDIETTESMRCGRHSI	257		
163	RTDSLWGLAERKRGWGYLFKNRPDGNAPNSFYRLPKIITQDIETTESMRCGRHSI	222		
258	QRHCRSETSKGYVCIQYDQKIVSGLRDNTIKIMQNTLECKRIILGHGYSVLCQYDE	317		
223	QRHCRSETSKGYVCIQYDQKIVSGLRDNTIKIMQNTLECKRIILGHGYSVLCQYDE	282		
318	RVIITGSSDSTVRVWDVNTGEMTLIHCEAVLHLFNNGMVATSCDKRSIAVWMAAP	377		
283	RVIITGSSDSTVRVWDVNTGEMTLIHCEAVLHLFNNGMVATSCDKRSIAVWMAAP	342		
378	TDITLRVLYGHRAAVNVVDEDKIYVSAAGDRITIKWNTSTGEFVTLNGHKGRLACQ	437		
343	TDITLRVLYGHRAAVNVVDEDKIYVSAAGDRITIKWNTSTGEFVTLNGHKGRLACQ	402		

100

A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-605 <BAD>
 A:Cross-references: EMBL:294864; PIDN:CAB08168.1; GSPDB:GN00066; SPDB:SPAC57A10.05c
 A:Experimental source: EMBL strain 972h; cosmid c57A10
 C:Genetics:
 A:Gene: SPDB:SPAC57A10.05c
 A:Map position: 1
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 19.5% Score 590.5; DB 2: Length 605;
 Best Local Similarity 29.0%; Pred. No. 8.7e-38;
 Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

```

QY 80 SSMIVKORRLKSAVSEKELCVKFEOWSE---SDQVEVEHLISOMCHYGHINSYL 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 46 SSM---HNEISGSEKSHQREVAWAFSEASCSERKALQILNCCSSLSLFASSLT 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 KPMLODFITLAPRGLDIAENILSYLDAKSLCAAEVCKEYRTSPGMLMKLIERM 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 DSLVRLDFISLPLV---BISFRILSFIDARSLCAQAAVSKHMKELADDVIMHRCCEH 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 197 VRTDSLIRGLAERRGQVLFKN----- 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 INRK-----CEKCGWGLPRLERNTLYAAKASTQKREYELTKRQVDAHESPVKAKLD 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 -----KPPDGNAPNSF-----YRALYKIIIDITETIESNRCGRHSLO 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 DVTSSNEETISSKPPSPNSDSKFLPKTRPKVEVARECR---VECNMRHGR----- 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 259 RIHCR-----SEMSKGYCQYDDQKIVSGLRDNTIKIMKNTLECKRILITGHTGSLCQ 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 ---CROVVLSGHSDGWCQQLVKNILASGSYDITRLMIAFQOVALLEGSSGVTCQ 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 YDERVITGSSDSTVRYMVDVNTGEMLNTLIHCEAVLHLRFNNGMVTCSDRSIAVMD 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 PDCKLISGSMKTRIKIMRTSECSIIIGHDSVLCITFSTLIVSSADCTVAKLMHF 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 ASPDTITLRLVLRVAGRAAVNVDF--DKYIVASAGRTIKYWNSTCEFEVTLNCHKRG 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 SGKRITLR---GHTGPNYSVRIIDRGVLVSGSDSTIKIWSLETNCLMFTFSAHIP 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 433 IACLOVRDLVYVSGSDNTIRLMDIFCGACLRVLEGEELVRCIRPDNRKIYSGAYDGI 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 VOSLADLSRLFCSLDGTIKWDIEKKKCVHTLFEGHIGVEIADHRLISGADGVV 496
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 KVMIDLVAALDPRAPACTLCLRLVESHGRVFLQDFEQIVSSSHDDTLIMDELNDP 550
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 KWEACE-----CVHTLKNHSEFVTSVALGDCEVSGSDGKILYLMENAP 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5
 S49932
 MET30 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein Y19905.02; protein Y11046w
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 26-May-2000
 C:Accession: S49932; S43750
 R:Odell, C.; Bowman, S.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S49931
 A:Accession: S49932
 A:Molecule type: DNA
 A:Residues: 1-640 <CODE>
 A:Cross-references: GB:247047; EMBL:248661; NID:6603997; PID:9763300; MIPS:Y11046w
 R:Thomas, D.; Cherest, H.; Barbey, R.; Surdin-Kerjan, Y.
 submitted to the EMBL Data Library, December 1993
 A:Reference number: S43750
 A:Accession: S43750
 A:Molecule type: DNA
 A:Residues: 1-60, '1', 62-640 <THO>
 A:Cross-references: EMBL:L26505; NID:g432493; PID:g432494
 C:Genetics:

A:Gene: SGD:MET30
 A:Cross-references: SGD:S0001308; MIPS:Y11046w
 A:Map position: 9L
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 F:298-339/Domain: WD repeat homology <WD1>
 F:338-369/Domain: WD repeat homology <WD2>
 F:374-409/Domain: WD repeat homology <WD3>
 F:417-450/Domain: WD repeat homology <WD4>

Query Match 18.0% Score 545; DB 2: Length 640;
 Best Local Similarity 28.9%; Pred. No. 3.2e-34;
 Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

```

QY 34 IPEKNSLRQTVNSCARCLN--QETVCLASTAMKTEVCAR--TKLANGTSSMIVKORRLS 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 LPEYNTKFCYRHNDDIOFSPHTACQKODLKRQEIINAKIPLQEOSDIIHIIKYS 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 ASYEKREKLCYKFEOWSESOVEVEVEHLISOMCHYGHINSYLRKMLQDFITLAPR 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 NSNDKIRKL-----ILDGILSTSCFPOLSYLSSLVTHMKIKIDFISLP-- 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 GLDHTAENILSYLDAKSLCAAEVCKEYRTSPGMLMKLIERMVRTDSLIRGLAERR- 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 --QELSLKILSTLDCQSLCNAFRCRKWKQKLADDRVYHMCQHT-----DRKC 236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 ---GWGYLFRNK-----PPDGNAPNSFYRALYKIIIDITETIESNRCGRHSLO 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 PNCWGGLPLHMKRRARIQNSTGSSNADITQTRTPKVIYRERFYESWVRG----- 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 259 RIHCRSEK-----GVYLOQYDDQKIVSGLRDNTIKIMKNTLECKRILITGHTGSLCQ 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 --HCHIOEFKGMVDVLRLQFNRYRLFGSYDSTGLWDLPTGRLIRLSGHSQGVITLY 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 YDERVITGSSDSTVRYMVDVNTGEMLNTLIHCEAVLHLRFNNGMVTCSDRSIAVMD 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 PDCKLITSLDKTRIKIMRTSECSIIIGHDSVLCITFSTLIVSSADCTVAKLMHF 409
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 ASPDTITLRLVLRVAGRAAVNVDFDQYI--VSAAGRTIKYWNSTCEFEVTLNCH- 429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 ESRTCYTLR---GHTGPNYSVRIIDRGVLVSGSDSTIKIWSLETNCLMFTFSAHIP 465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 430 KRGIACLOVRD--RLVSGSS----- 448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 VOKIIPLIKVDENLATDNTSDGSSPODDPTMTGADSDSPNSBOETVDENIPYTHL 525
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 449 -----DNTIRLMDIFCGACLRVLEGEELVRCIRPDNRKIYSGAYDGIKWDVLAALD 503
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 526 LSCGIDNTIKIMDVYTGCIKIRQFGHVGCVWDIAADNFRITSGSHDSIKYMDLQSC--- 582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 504 RAPACTLCLRLVESHGRVFLQDFEQIVSSSHDDTI 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 583 -----KCMHFP---NGR--RLQRETOHTQTOGSLGDXV 609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6
 T6660
 sulfur controller-2 protein [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
 C:Accession: T6660
 R:Kumar, A.; Paletta, J.V.
 Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995
 A:Title: The sulfur controller-2 negative regulatory gene of Neurospora crassa encode
 A:Reference number: T6660
 A:Accession: T6660
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-650 <KUD>
 A:Cross-references: EMBL:U17251; NID:9806757; PIDN:AAA68968.1; PID:9806758
 C:Genetics:
 A:Gene: scon-2
 A:Map position: 3
 A:Introns: 75/3; 319/1; 354/1

OY 460 GACIARVLEHHEELVRCIRRDNRKIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTLVHEH- 518
 DB 180 GKLLRMLGHEHDLIRTVRNDKIRIVSGYDGTVRIN-----FNTGEQHC 224
 OY 519 -----SCRFRRLQFDEFOIVSSSHDDTLILIMDF 546
 DB 225 VLHNSRNSRVFGIQLPHRRTIACHTSSELVWNE 258

RESULT 9

T18521

beta transducin-like protein - Podospora anserina

C:Species: Podospora anserina

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18521

R:Saupé, S.; Turcq, B.; Bequeret, J.

Gene 162, 135-139, 1995

A:Title: A gene responsible for vegetative incompatibility in the fungus Podospora anser

A:Reference number: Z18944; MUID:96009891; PMID:7557402

A:Accession: T18521

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1356 <SAU>

A:Cross-references: EMBL:L28125; NID:g607002; PID:g607003; PIDN:AAA85775.1

C:Genetics:

A:Gene: het-el

A:Insertions: 761/3

Query Match 15.0%; Score 455.5; DB 2; Length 1356;
 Best Local Similarity 26.0%; Pred. No. 8.2e-27;
 Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

OY 28 EPRKIIIPKNSLRQITNSCAR-LCLNQETVCLASTAMK-----TENCVAKTKLANGTSS 81
 DB 581 EMPDVSDDESELEIVKLCGSPFLIRKTVYVHOSAKPFLIGTASDKASNKASQDAFE 640
 OY 82 MIYPR-----QRKLSAYEKEK--LCVYFPEQMSDOVEFEH--ISOMCHY 127
 DB 641 LVFTGTIEDVYIIFRSLVWMSGKLARDIYCLNAPFLIDNVRPPDPLATVRYSCY 700
 OY 128 QHGHINSYLRP-----MLQRD-----FTT-----ALPANGLDHI--AEN 159
 DB 701 WIDHLRDLVSTSSKVVHLLDDDDIRHPLTLYLWLEALSLRLAP--EGINAIROLES 759
 OY 160 ILTSLDAKSLCAALVCKEWRVTSDDAMKILIER--WVRDTSRLGLAERGMGOYL 216
 DB 760 LIGHITIRGLTA--IVRDGIRF--ALSTRMILEKAPLQATYSAL-----V 800
 OY 217 FKNRPDGNAPNSFYALY----PKTIODIETIESNMRCGRSL----- 257
 DB 801 F-----APTDSMIKKIFKEKPGMISTISVVEAEINACTQTLLEGHSSVLSVAFSAD 852
 OY 258 -QRHCHNS-----ETSKG-----YCCLOY--DDQKIVSGLRDNITKIMDK 294
 DB 853 GORVAGSGDDTKIKIMDTASGTQTLLEGHGSVWVAFSFDRERVAGSGDDTKIKIMDA 912
 OY 295 NTLCKRLITGHGTSVLCLOY--DERVITGSSDSTVYVMDVNTGEMNTLIHICEAVLH 352
 DB 913 ASGTCTQTLLEGHGSVWVAFSFDRERVAGSGDDTKIKIMDAAGTCTQTLLEGHSSVLS 972
 OY 353 LRFN-NGMMVTC-SKDRSIAVMDASPTDITLRLVYLGHRAAVNVDF--DDKIVASG 408
 DB 973 VAFSPDQORVAGSGDDTKIKIMDTASG--TCTQTLLEGHGSVWVAFSFDRERVAGSGD 1029
 OY 409 DRTIKVWNTSCEFRVRLTNGHKRGITACLOYR--DRLVYSSDNTIRLMDIECGACLRVL 466
 DB 1030 DRTIKIMDTASGTCTQTLLEGHGSVWVAFSFDRERVAGSGDDTKIKIMDAVSGTCTQL 1069
 OY 467 EGHEELVRCIRF--DNKRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTLVHSGRVER 524
 DB 1090 EGHGDSVWVAFSFDRERVAGSGDDTKIKIMD-----AASGT--CTQTLLEGHGSVW 1140

OY 525 LQF--DEFOIVSSSHDDTLILIMDLNDPAAQ 553
 DB 1141 VAFSPDQORVAGSGDDTKIKIMDAAGTCTQ 1171

RESULT 10

T45136

WD repeat protein popl [imported] - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T45136; T40157

R:Komitani, K.; Toda, T.

submitted to the EMBL Data Library, September 1996

A:Description: Fission yeast WD repeat protein Popl is involved in maintenance of plo

A:Reference number: Z2925

A:Accession: T45136

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-775 <KOM>

A:Cross-references: EMBL:Y08391; PIDN:CAA69671.1

A:Experimental source: strain h-972

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.

submitted to the EMBL Data Library, March 1998

A:Reference number: Z21842

A:Accession: T40157

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-775 <WOOD>

A:Cross-references: EMBL:AL022103; PIDN:CAA17898.1; GSPDB:GN00067; SPDB:SPBC262.18

A:Experimental source: strain 972h; cosmid c262

C:Genetics:

A:Gene: SPBC262.18

A:Map position: 2

Query Match 14.9%; Score 453; DB 2; Length 775;
 Best Local Similarity 25.8%; Pred. No. 5.8e-27;
 Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;

OY 9 QEKALFNMSSEREDC-----NNGEPRKIIIPKNSLRQTY--NSCARLCLNQET 56
 DB 121 REKCLKRRNSLSSNLHANKRFLFNSQSGGNKKNTEFPSTNYSNFFYRNCDKSEVASET 180
 OY 57 VC-----LASTAMKTENCVAKTKLANGTSSMIYPRKOKLEA----- 92
 DB 181 TFSLDAPNNNSVNTSPSPMLGNDSTKRSQFPFPHSSSSHNSHEPVIYDFSSNPISHP 240
 OY 93 -----SYEKEKEKEL-----CVKYFPEQMSDOVEFEH--ISOMCHYQHGHINSYLRKPLQDRF 144
 DB 241 SNHLSOKNAVLALAOISSFEKLPESVROYILFHLISCKGKHAVQNIHKLILPIFOKNF 300
 OY 145 ITALPARGLDHIENILSYLDAKSLCAELVCKEWR--VTSQGMIMKTL-----I 193
 DB 301 LTGFPA-----ETNNLVTLHDAPSLCAVSGVSHHWKLVSSNBEIMKLSFLKDGCFWDSI 356
 OY 194 ERVAVRDSILMRGLAERGMGOYLFKNKPPDGNAPNSFYALYPKTIODIETIESNMRCG 253
 DB 357 DSKIRMLCLOSLA-----CAIKKRYFPHFNLREK----- 389
 OY 254 RSHLQRIHCHSETSKGYVC---LOYDOKIVSGLRDNITKIMDKNTLCKRLITGHTGVS 310
 DB 390 LHAPEKIKRSPFIHGVRILTKLOPDDKILIVTSGPRINITYDTKGVILRSLEHEGIV 449
 OY 311 LCIQYDERVITGSSDSTVYVMDVNTGEMNTLIHICEAVLH-LRFNNGMMVTC-SKDRSIA 370
 DB 450 WTEFEYDVTLVGTSTDRVWMLRTGCECKQVEYGHSTIRCIKIVQGNOSTDTD----- 505
 OY 371 VMDASPTDITLRLVYLGHRAAVNVDFDDKIVASAGDRTIKVWN----- 416
 DB 506 -----DVEKE-----NRPASNDANSMPPIIISRRDCTILMSLPCLDPPFVNVNE 552
 OY 417 -----TSTCEFRVRLTNGHK--RGITACLOYRDLVWSSDNTIRLMDIECGAC 462

Db 553 NPDQNDFTSATNPDEYIRLGRHTDSREYACU--GDLIVASXGTLRWKASTGVC 609
 QY 463 LRVLEGHELVRCIRFENKR--IYSGAYDGKIKWMDVLAALDPRAPGTLCLRTIVHSG 520
 Db 610 LHVLRGHWGRVYSTVINSRQOCISAGTDKIRIMWL-----ESGEL-IQTLHGSHN 660
 QY 521 RVFRLOFDEFOIVSS--HDDTILIMD 545
 Db 661 LVSVTFNONTIVASAPPTSLRWMD 687

RESULT 11

MD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AE1810
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AE1810
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1227 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA077553.1; PID:417135007; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 A:Genetics: alr0029

Query Match 13.6%; Score 413.5; DB 2; Length 1227;
 Best Local Similarity 23.9%; Pred. NO. 1.3e-23;
 Matches 140; Conservative 114; Mismatches 208; Indels 125; Gaps 22;

QY 10 EKALKPMNSERE-----DCNNGEPKRIITEKNSLRQTYSCARLCINDFTVCLA 60
 Db 378 EKVYVMSLGANNRYVSFRLEKQDWLTTESPIKVMALSLR-----RSLEIKASPLLI 431
 QY 61 STAKTENCVAKTKLANGTSSMIVPRKRLSAYEKELCVKYPEQSESDOYEVNHL 120
 Db 432 EKASSYQGEKEASSKFGLESYVM--EYITAKFIENS-----LEHSQKRIKDF-- 478
 QY 121 ISCMYCHYGHINSYKPMLODFITLAPARGLDHIAENILSYDAKSLCAAEIVCREWY 180
 Db 479 -----INTY--PLMK-----ARSLDYI----- 493
 QY 181 RYTSQMLMKKLIEMVVRDLSMRGLABRGQVLFKNKPPDGAPN--SYRRLYRK 238
 Db 494 RQIQRLLIEPYKQKLLNIFGELEHLRRLMGLTQKEPLFKGVAAGNLINLLRQLDL 553
 QY 239 ITQDIETISN-----WRCGRHSI--QRTHRS-----ETSKGVYCLQY--D 276
 Db 554 KLPDSPPDLSGRDPSGLTIWQAYFEVKLKEITIFANSDLGVSFTLEMSSVSVSKRSPD 613
 QY 277 DQKIVSGLENDIKITMDNTLECKRLLGHGVSVCLOY--DEVYITIGSSDSTVRWVY 334
 Db 614 GYFPAFGMLNGEIRLMQTSNKLQKRTYKGTAWMAFAFSPDSMLASGSADSTIKIMDV 673
 QY 335 NTGEMINLTIHRCVAVLHRR--NNGMVAVTSKSRSTAVWMAAPPTDITLRVLYVGRRA 392
 Db 674 HNGECLTKLSKNTKNSVAVSPGRILASQDQITIKIMDIANG--NCOQTILGHDDW 730
 QY 393 VAVVDF-----DDK--YIYASAGDRTIKYKWTSTCEVFTLNGHRGIALQYR--DRLVY 444
 Db 731 VWSVTFSPVTDRLPLLAASSADQIKIMDVATGCKLTKGHTREHVSFSFSDQTLA 790
 QY 445 SSSSDNTIRLMDIEGACLRVLEGEHELVRCIRF--DNKRIVASAYGKTKVMDVLAALD 502
 Db 791 SSGSDSTYRLMDVKTGCQWQIFEGHKKYIVSVRSPDQGLTASGGERSTIKIMDIORG-- 848
 QY 503 PRAAPAGTLCRLTIVHSGRVRLQF--DEFOIVSSSHDDTILIMDPL 547

Db 849 -----ECVWTLHGSHSQWAIASFDPGRLLISCSDDQTARLMVDYI 888

RESULT 12

cell division control protein CDC4 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YPL009W
 C:Species: Saccharomyces cerevisiae
 C:Date: 02-Sep-1995 #sequence_revision 12-Apr-1996 #text_change 26-May-2000
 C:Accession: S56245; S48310; A26867; S62304
 R:Murakami, Y.; Maitou, M.; Hagihara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sas
 submitted to the EMBL Data Library, May 1995
 A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces
 A:Reference number: S56245
 A:Accession: S56245
 A:Molecule type: DNA
 A:Residues: 1-779 <MUR>
 A:Cross-references: EMBL:D50617; NID:g836685; PIDN:BA09229.1; PID:d1009870; PID:g836
 R:Churcher, C.
 submitted to the EMBL Data Library, September 1994
 A:Reference number: S48310
 A:Accession: S48310
 A:Molecule type: DNA
 A:Residues: 1-579 <CHU>
 A:Cross-references: EMBL:246255; NID:g559925; PIDN:CAA6341.1; PID:g559926; MIPS:YPL0
 R:Jochem, U.; Byers, B.
 J. Mol. Biol. 195, 233-245, 1987
 A:Title: Structural comparison of the yeast cell division cycle gene CDC4 and a relat
 A:Reference number: A26867; MUID:88011240; PMID:3309335
 A:Accession: A26867
 A:Molecule type: DNA
 A:Residues: 1-459, 'E', 461-779 <YOC>
 A:Cross-references: EMBL:X05625; NID:g3502; PIDN:CAA29113.1; PID:g3503
 R:Murakami, Y.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S62302
 A:Accession: S62304
 A:Molecule type: DNA
 A:Residues: 1-779 <MUR>
 A:Cross-references: EMBL:D31600; NID:g836814; PIDN:BA06495.1; PID:d1007066; PID:g836
 C:Genetics:
 A:Gene: SGD: CDC4
 A:Cross-references: SGD:S0001885; MIPS:YPL009W
 A:Map position: 6L
 C:Function:
 A:Description: Initiation of DNA replication; separation of the spindle pole bodies t
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 C:Keywords: cell cycle control
 F:459-494/Domain: WD repeat homology <WD1>
 F:628-659/Domain: WD repeat homology <WD2>

Query Match 13.2%; Score 399; DB 2; Length 779;
 Best Local Similarity 24.3%; Pred. NO. 9.1e-23;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 IYCLASTAKTENCVAKTKLANGTSSMIVPRKRLSAYEKEL-----ELCVKFEQW 108
 Db 194 TTPPLAKTKITNN-----NNNADLIESKDSITISPEYISDITFSAINNNLPHAYK-- 244
 QY 109 SESDOVEFEHLISOMCHYGHINSYKPMLODFITLAPARGLDHIAENILSYDAS 168
 Db 245 -----NLFRLVANMRSELSDLGITLKNLKRDLTSLP-----EISLITPNYIQFED 294
 QY 169 ICAAEVLCKEYR-VTSDGMLMKLI--EMVVRDLSMRGLABRGQVLFKNKPPDG 225
 Db 295 IINSLGVSONNKKITRKSTSLMKRLISENFV-----SPKGF 331
 QY 226 APPNSFYRALYKTIQD-----IEF--IESNMRCGRHSIORHCRSFTSGVYCLQYD 276
 Db 332 NSLNLKLSQKYPKLSQGDRLRLSLLENIFLKNWYKPYQPTITLKGMTSVYICLOPE 391
 QY 277 DQKIVSGLRDNTIKITMDNTLECKRLLTGHGVSVCLOYDE-RVITIGSSDSTVRWMDV 335

DB 392 DNVTITGADGKMTKRTYDSINKKFLQLSCHDGVWALKAHGGIILVSGSTDRTRVWMDIK 451
QY 336 TGEMLTLIHCEAVLHLRFNNMGWTCCKRSIAVWMASTPDTITLRVLVGHRAAVV 395
DB 452 KG-----CCT-----HVFQGHNSTVAC 468
QY 396 VDFDD-----KIYVASAGDRTIKVWMT-----STCE-----PVRTLNG 428
DB 469 LDIVETKNTIKYITGSRDWTLLHWKLPKESVDPHGEEDHPLVFTPTPEENPFVGVLRG 528
QY 429 HKKGACIQLYRDLVYVSGSSDNTIRLMDIEGACILVRLVGHHELVLCIFEDN--KRIYVG 486
DB 529 HMASTVTSRGKNTIVSGSTDNLTIVMDVAKMCLTILSGHTRDYSTIYDHRKRCISA 588
QY 487 AYDGKIKVWL-----VALDPRAPAGTL--CLRTLVHSGRVRLQDFEFOIYSS 535
DB 589 SMDTIRIMDLNINWNGEVSATNSAPCAKILGAMTYLQGHVTLVGLRLSDKRLVSA 648
QY 536 SHDITLIDFLNDPAAQAEPSPSRITTY 566
DB 649 AADGIRGWD-AND-----YSRKFSY 668

RESULT 13

F-box/WD-repeat protein pop2 - fission yeast (Schizosaccharomyces pombe)

N.Alternate names: proteolysis factor sud1p

C.Species: Schizosaccharomyces pombe

C.Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C.Accession: T43557; T38794; T43798

R.Wolf, D.A.; Jackson, P.K.

submitted to the EMBL Data Library, December 1997

A.Description: Fission yeast pop2 encodes a novel F-box/WD-repeat protein involved in th

A.Reference number: 222576

A.Accession: T43557

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-703 <MW>

A:Cross-references: EMBL:AF038867; PIDN:AMB95480.1

A:Experimental source: strain h-972

R.Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1996

A:Reference number: 221812

A.Accession: T38794

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Residues: 1-703 <GEN>

A:Molecule type: DNA

A:Cross-references: EMBL:Z98602; PIDN:CA811275.1; GSPDB:GN00066; SPDB:SPACAD7.03

A:Experimental source: strain 972h-; cosmid c4D7

R.Jallepalli, P.V.; Tien, D.; Kelly, T.J.

Proc. Natl. Acad. Sci. U.S.A. 95, 8159-8164, 1998

A>Title: Sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Runt proteins fo

A:Reference number: 222686; MUID:98318628; PMID:9653157

A.Accession: T43798

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-703 <CAL>

A:Cross-references: EMBL:AF064515; NID:g3293382; PIDN:AMC39496.1; PID:g3293383

C:Genetics:

A:Gene: pop2; sud1; SPACAD7.03

A:Map position: 1

C:Function:

A:Description: regulated to prevent spontaneous re-replication

Query Match 13.1%; Score 396.5; DB 2; Length 703;
Best Local Similarity 23.3%; Pred. No. 1.2e-22;

Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

QY 14 KFNNSERED--CNGGEPKRI-----IPEKNSLRQTYNSCARLCT---NOETVCL 59
DB 96 KFMESTIENSYKCLDHPSPVPGDPSISLVPQRNPL-VSHSLPRKISIDRNNIKID 154

QY 60 ASTANKTENCYAKTRKLANGTSSMIVPKOKRSASYEK-EKELCYKEFEQMSQVEFE 118
DB 155 NSISSNSDNPSPPKV--DTSNTVSPGSKPISELEDNINOSIYOTEDPEGIG-SYAF 211
QY 119 HLISOMCHYOHGH-INSYLRKPLQDFITLAPAGLDHIAINIISYDAKSLCAELVCK 177
DB 212 FOLLRSCNRQSMRLNLNCEPLKADILSNLP-----SIVSILNLDIHSLCRLVSP 267
QY 178 EMYRV-----TSDGMMLKRLIERMVRTDSLMRGLAERRGMGYLFRKPPDNAPP-----N 229
DB 268 TWRNRLDVHTS---YKIMAF-----SLFGQINENMW-----KYANPNLRPPFLAND 312
QY 220 SFYRALYRKTIIDTLETTESNNRCGRHSIQR----- 259
DB 313 QISDDYFPEIRK-----RHFLNRRKWLFPSPISPHLSFPIHPNMTISLLH 360
QY 260 -----IHCRSETSK-GVYCYLDYDQKIVSGIRNDITKMDNT 296
DB 361 KDRITITGSGTIOIHNAITGVLEARKLKGKGVAVAHENTLVSGSIDTVRWNIER 420
QY 297 LECKRILTGHTGSVLCIOLY-----DERVITGSSDSTVRWVDNTGEML 340
DB 421 AKCTHIFRGHISILRCLELIVPSRLIRHGVETVEPDQPIYVSGSRDHTLRW----- 472
QY 341 NTLIHCEAVLHLRFNNMGWTCCKRSIAVWMASTPDTITLRVLVGHRAAVVDFPD 400
DB 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNPFV--HTLIGHTDSVTRISYGY 517
QY 401 KIYVASGDRITKIVNTSTCEPFTLNGHKGICLOYRD--RLVSGSSNTRIRMDIE 458
DB 518 DLIVSGSTDSIRIRVSTGECLYHLRSHLRITSVLPEPRNRCISGSMKSTRVMDLS 577
QY 459 CGACLRVLEGEHELVRCLRFNPKRIVSGAYDGKIKVMDLVAALDPRAPAGTICRTVHEH 518
DB 578 TGTCTVYLEGHDAFVTLTNVFNRLISGSADSTIRIMD-----LMTGR-----LMVLPN 628
QY 519 SCRVRLODFEQIYSSSHDITILMD 545
DB 629 SCYISFVSDEHKIT-SSNDGVSVKLMD 654

RESULT 14

WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120) plasmid PCC7120alpha

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C.Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C.Accession: AT2493

R.Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:2159285; PMID:11759640

A.Accession: AT2493

A>Status: preliminary

A:Molecule type: DNA

A.Residues: 1-1189 <NR>

A:Cross-references: GB:BA000020; PIDN:BA878213.1; PID:917135667; GSPDB:GN00180

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alt7129

A:Genome: plasmid

Query Match 12.5%; Score 378.5; DB 2; Length 1189;
Best Local Similarity 26.7%; Pred. No. 6.5e-21;
Matches 96; Conservative 63; Mismatches 107; Indels 93; Gaps 10;

QY 280 IYSGLRDNTIKIMDKNTLECKRIILTGHTGSVLCIOLYD--ENVITGSSDSTVRWVDNTG 337
DB 745 IASGSFDONVKKLMDIHTGKCVMTLOGHGVYTSVAFNPKMDLLSGSYDQSVKWDKRTG 804
QY 338 EMLNLTILHCEAVLHLRFN----- 356

Db 805 RCLDPLTKHTRNIRWSVAHFPOGHLFVSGDDHAAKIMWELGTQCCKTFQGHSNATYTHA 864
QY 357 ---NMMWTCCKDRSIANWD--MASPTDILR-----RYLVGHAAYNVNDEDD--KIYVS 405
Db 865 NMESHLASGHEDOTIKLMDLNLHSPKRSNVTHPEFRLQGHNSRNFVSFSTGQLLAS 924
QY 406 ASGDRTIKVMWNTSTCEFRVTLNGHKGRIACLOVR--DRLVYSGSSDNTIRLMDIEGACL 463
Db 925 GSADRTIKLWSPHTGQCLHTJHGHWSWMAIAFSLDDKLASSYDHTVKIMDVSSGQCL 984
QY 464 RYLBGHELVLCIRF--DNKRIYSGAYDGIKIMDL-----VAALDP 503
Db 985 QTLQGHPEGYLAFAFSCDGKTLFSSGKELVKQWDVETGYCLQTEADSNRWMAVAVSRL 1044
QY 504 RAPAGT-----LCRLTVEHSGRYERLQF--DEFOIVSSSHDITILMD 545
Db 1045 NQYLATGGDDSVRLMDIGKGVCTRTSGHTSOYICILFTKDGRRMISSSSDRTIKIMN 1103

RESULT 15

AC1842

MD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AC1842

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001

A:title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC1842

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1747 <CUR>

A:Cross-references: GB:BA000019; PIDN:BAF7807.1; PID:q17135261; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0283

Query Match

Best Local Similarity 32.7%; Pred. No. 2.1e-20;

Matches 84; Conservative 50; Mismatches 107; Indels 16; Gaps 7;

QY 276 DDOKIVSGLRDNTIKIMDKNTLECKRIITLGHGTVLCLQY--DERVITGSSDSTVVRWD 333
Db 1487 DGTIATASADNTIKLMDSCQQLIKTLTGHKRITLTFHPDNOQTIASGSADKTIKIR 1546
QY 334 VNTGEMNTLTHHCEAVLHLRF--NNGMWTCSKDRSIAYMDASPTDITLRRYLVGHR 391
Db 1547 VNDGQLLRTLTGHNDEVTYVNSPQGLASGSTDNTVKIMQ---TDGRLINKITGHGL 1602
QY 392 AVNVVDF--DDKYIVSASGRTIKVMNTSTCEFRVTLNGHKGRIACLOVR--DRLVYSGS 447
Db 1603 AIAVSKFSPDSHTLASASMDNTIKLMQVTDOKLINLNGHIDGVTSLSPDGEIILASGS 1662
QY 448 SDNTIRLMDIEGACLVLESGHELVLCIRF--DNKRIYSGAYDGIKIMDLVAALDPRA 505
Db 1663 ADNTIKIMNLPNATLTKLGHPEKINTLAFSPDGKTLISGGEADAGVMVNL--DLDDL 1720
QY 506 PAGTLCRLTVEHSGRY 522
Db 1721 QQGGCDRLTDYLOHNSNV 1737

Search completed: February 20, 2003, 09:55:48
Job time : 28 secs

- 1: genEmb1: *
- 2: gb_ba: *
- 3: gb_hlg: *
- 4: gb_in: *
- 5: gb_om: *
- 6: gb_ov: *
- 7: gb_pal: *
- 8: gb_ph: *
- 9: gb_pl: *
- 10: gb_pr: *
- 11: gb_ro: *
- 12: gb_scs: *
- 13: gb_sy: *
- 14: gb_un: *
- 15: gb_vl: *
- 16: em_ba: *
- 17: em_hum: *
- 18: em_in: *
- 19: em_mu: *
- 20: em_om: *
- 21: em_or: *
- 22: em_ov: *
- 23: em_pat: *
- 24: em_ph: *
- 25: em_pl: *
- 26: em_ro: *
- 27: em_scs: *
- 28: em_un: *
- 29: em_vl: *
- 30: em_hlg_hum: *
- 31: em_hlg_inh: *
- 32: em_hlg_other: *
- 33: em_hlg_mus: *
- 34: em_hlg_pin: *
- 35: em_hlg_pin: *
- 36: em_hlg_mam: *
- 37: em_hlg_vrt: *
- 38: em_sy: *
- 39: em_hqo_hum: *
- 40: em_hqo_mus: *
- 41: em_hqo_other: *

Result No.	Score	Query Match	Length	DB	ID	Description
1	2151	100.0	2151	6	AX019507	AX019507 Sequence
2	2151	100.0	2151	6	AX057166	AX057166 Sequence
3	2151	100.0	2151	9	AF129530	AF129530 Homo sapi
4	2151	100.0	2151	9	HSBTRCP	Y14153 Homo sapien
5	2030.4	94.4	2308	9	BC027994	BC027994 Homo sapi
6	1682.2	77.3	2175	10	AF081887	AF081887 Mus muscu
7	1653.2	77.3	1818	9	AF101788	AF101788 Homo sapi
8	1630.4	75.8	2081	10	BC003969	BC003969 Mus muscu
9	1567.2	72.9	1979	10	AF110336	AF110336 Mus muscu
10	1451.2	67.5	1712	10	AF099932	AF099932 Mus muscu
11	1442.8	67.1	1710	10	AF112979	AF112979 Mus muscu
12	1017.8	47.3	1671	5	XELSTRCP	M98268 African cla
13	883.4	41.1	2134	5	AB033279	AB033279 Homo sapi
14	883.4	41.1	2252	9	AB033280	AB033280 Homo sapi
15	883.4	41.1	2274	9	AB033281	AB033281 Homo sapi
16	883.4	41.1	4230	9	AB014596	AB014596 Homo sapi
17	883.4	41.1	4395	9	BC026213	BC026213 Homo sapi
18	882.4	41.0	1653	9	AF176022	AF176022 Homo sapi
19	877.8	40.8	4030	10	BC034261	BC034261 Mus muscu
20	853.2	39.7	1970	10	AY038079	AY038079 Mus sapi
21	757.8	35.2	67008	9	AP000252	AP000252 Homo sapi
22	757.8	35.2	100000	9	AP000012	AP000012 Homo sapi
23	757.8	35.2	100000	17	AP0000134	AP0000134 Homo sapi
24	757.8	35.2	149298	9	AP0000031	AP0000031 Homo sapi
25	757.8	35.2	340000	9	AP001711	AP001711 Homo sapi
26	716.4	31.4	2534	3	AT118898	AT118898 Drosophill
27	673.2	31.3	2154	3	AF032878	AF032878 Drosophill
28	673.2	31.3	2367	3	AF222924	AF222924 Drosophill
29	673.2	31.3	3532	3	AF222923	AF222923 Drosophill
30	583.2	27.1	188524	2	AC112602	AC112602 Rattus no
31	518.8	24.1	1561	6	AX057148	AX057148 Sequence
32	512	23.8	934	5	XJ063921	XJ063921 Xenopus lae
33	502.6	23.4	1827	5	XJ063922	XJ063922 Xenopus lae
34	463.6	21.6	2983	10	BC0080552	BC0080552 Mus muscu
35	454	21.1	1443	3	AF339101	AF339101 Heteroder
36	453.2	21.1	2700	3	AF275523	AF275523 Caenorhab
37	333.2	15.5	108503	9	AL627424	AL627424 Human DNA
38	288	13.4	173634	3	AC009344	AC009344 Drosophill
39	277.2	12.9	430266	3	AC003733	AC003733 Drosophill
40	277.2	12.9	40472	2	AC014085	AC014085 Drosophill
41	252.8	11.8	17344	3	U28730	U28730 Caenorhabdi
42	232.8	10.8	143078	9	AL454563	AL454563 Human DNA
43	225.4	10.5	231703	2	AC126454	AC126454 Mus muscu
44	225	10.5	266	10	FS91178509	FS91178509 Mus muscu
45	216.2	10.1	261252	2	AC111315	AC111315 Rattus no

ALIGNMENTS

RESULT 1				
AX019507				
LOCUS	AX019507	2151 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO9338969.			PAT 07-SEP-2000
ACCESSION	AX019507			
VERSION	AX019507.1	GI:10043427		
KEYWORDS	.			
SOURCE	synthetic construct.			
ORGANISM	synthetic construct.			
REFERENCE	artificial sequences.			
AUTHORS	1 (bases 1 to 2151)			
TITLE	Areznani,S.F., Concordet,J.P., Kroll,M., Durand,H., Benarous,R. and Margotin,F.			
JOURNAL	Protein humane beta -tcp			
	Patent: WO 9338969-A 1 05-AUG-1999;			

ARENZANA SEISEDOS FERNANDO (FR); CONCORDET JEAN PAUL (FR); INST
NAT SANTE RECH MED (FR); KROLL MATHIAS (FR); DURAND HERVE (FR);
BENNAOUS RICHARD (FR); MARGOTIN FLORENCE (FR); PASTEUR INSTITUT
(FR)

FEATURES
source Location/Qualifiers

1..2151
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="ADNC codant pour la proteine humaine BTCP"
70..1779
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07740.1"
/db_xref="GI:10043428"

CDS

BASE COUNT 628 a 467 c 513 g 543 t
ORIGIN
Query Match 100.0%; Score 2151; DB 6; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCCTGTCGCTGGCGCTGGACCAAGGGGGCGCCCGCGGAGAGCGGACCGAGTGGCC 60
Db 1 TCCTGTCGCTGGCGCTGGACCAAGGGGGCGCCCGCGGAGAGCGGACCGAGTGGCC 60
QY 61 TCGGGGATTATGACCCCGCGGCGGCTGCTGCAAGAAAGGCACTCAAGTTATGAA 120
Db 61 TCGGGGATTATGACCCCGCGGCGGCTGCTGCAAGAAAGGCACTCAAGTTATGAA 120
QY 121 TCCTAGAGAGAGAGACTGTAAATATGGGCAACCCCTAGGAAGTAAATCCAGAGAG 180
Db 121 TCCTAGAGAGAGAGACTGTAAATATGGGCAACCCCTAGGAAGTAAATCCAGAGAG 180
QY 181 AATTCACCTTAAAGACACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
Db 181 AATTCACCTTAAAGACACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
QY 241 TGTTAGCAAGCAGCTGATGAAGACATGAAATGTGTGGCCAAACAACTTGGCAAT 300
Db 241 TGTTAGCAAGCAGCTGATGAAGACATGAAATGTGTGGCCAAACAACTTGGCAAT 300
QY 301 GGCACCTTCACATGATGTGTGCCAAGCAAGAAACTCTAGCAAGCTATGAAAGGAA 360
Db 301 GGCACCTTCACATGATGTGTGCCAAGCAAGAAACTCTAGCAAGCTATGAAAGGAA 360
QY 361 AAGGAACCTGTGTCAAACTATTGAGCAGTGTGCAAGTCAAGTCAAGTGAATTTGTG 420
Db 361 AAGGAACCTGTGTCAAACTATTGAGCAGTGTGCAAGTCAAGTCAAGTGAATTTGTG 420
QY 421 GAACATCTTATATCCCAAAATGTGCTATACCAACATGAGGCAATAACTGTAATCTTAAA 480
Db 421 GAACATCTTATATCCCAAAATGTGCTATACCAACATGAGGCAATAACTGTAATCTTAAA 480
QY 481 CCTATGTTGAGAGAGATTCAATACGTCTGTGCACGCTGGGGATTGATCATATGCT 540
Db 481 CCTATGTTGAGAGAGATTCAATACGTCTGTGCACGCTGGGGATTGATCATATGCT 540
QY 541 GAGAACATTTGTCATACGCGGATGCAATCACTATGCTGCTGTAACCTTGTGTGAAG 600
Db 541 GAGAACATTTGTCATACGCGGATGCAATCACTATGCTGCTGTAACCTTGTGTGAAG 600

QY 601 GAATGTACCGAGTAGACCTCTGATGGCAGTGTGTGGAAGAGCTTATCGAGAGATGTC 660
Db 601 GAATGTACCGAGTAGACCTCTGATGGCAGTGTGTGGAAGAGCTTATCGAGAGATGTC 660
QY 661 AGGACAGATTTCTGTGTGAGAGGCGCTGGCAGACAGCAAGAGAGATGGGACATTTATTC 720
Db 661 AGGACAGATTTCTGTGTGAGAGGCGCTGGCAGACAGCAAGAGAGATGGGACATTTATTC 720
QY 721 AAAAACAACCTCTCGAGAGGGAATGCTCTCCCACTCTTTTATAGAGACTTATTC 780
Db 721 AAAAACAACCTCTCGAGAGGGAATGCTCTCCCACTCTTTTATAGAGACTTATTC 780
QY 781 AAAATTTATCAAGCATTTGAGACATATGAATCTAATTTGAGATGTGGAACATAGTTTA 840
Db 781 AAAATTTATCAAGCATTTGAGACATATGAATCTAATTTGAGATGTGGAACATAGTTTA 840
QY 841 CAGAGAAATTCCTCCGAGAGTGAACCAAGCAAGAGAGACTTACGTTTACAGTATGAT 900
Db 841 CAGAGAAATTCCTCCGAGAGTGAACCAAGCAAGAGAGACTTACGTTTACAGTATGAT 900
QY 901 CAGAAATATGTAAGCGGCTTCGAGACACACATCAAGATCTGGGATATAAACAATTC 960
Db 901 CAGAAATATGTAAGCGGCTTCGAGACACACATCAAGATCTGGGATATAAACAATTC 960
QY 961 GAATGCAAGGGAATTCACAGGCGCATACAGTTCAGTCTCTGCTCCAGATATGATGAG 1020
Db 961 GAATGCAAGGGAATTCACAGGCGCATACAGTTCAGTCTCTGCTCCAGATATGATGAG 1020
QY 1021 AGAGTATCATTAACAGAGATCATGAGATTCACAGTTCAGAGTGGGATGAAATACAGGT 1080
Db 1021 AGAGTATCATTAACAGAGATCATGAGATTCACAGTTCAGAGTGGGATGAAATACAGGT 1080
QY 1081 GAAATGTAAACACGTTGATTCACATTTGGAAGCAGTTCTGCACTTGGCTTCAATAT 1140
Db 1081 GAAATGTAAACACGTTGATTCACATTTGGAAGCAGTTCTGCACTTGGCTTCAATAT 1140
QY 1141 GGCATGATGTGACCTGCTCCCAAGATCGTTCATGCTGTGATGGATATGGCTCCCA 1200
Db 1141 GGCATGATGTGACCTGCTCCCAAGATCGTTCATGCTGTGATGGATATGGCTCCCA 1200
QY 1201 ACTGACATTAACCTCCGAGAGGCTGTGTCGACACCGAGCTGCTCAATGTTGATGAC 1260
Db 1201 ACTGACATTAACCTCCGAGAGGCTGTGTCGACACCGAGCTGCTCAATGTTGATGAC 1260
QY 1261 TTTGATGACAGATGATGTTTCTGATCTGGGAGATAGAACTATTAAGATATGAAACACA 1320
Db 1261 TTTGATGACAGATGATGTTTCTGATCTGGGAGATAGAACTATTAAGATATGAAACACA 1320
QY 1321 AGTACTGTGAATTTGTAGAGACCTTAATGGAACAAAGAGGACTTCCGTTTGGAG 1380
Db 1321 AGTACTGTGAATTTGTAGAGACCTTAATGGAACAAAGAGGACTTCCGTTTGGAG 1380
QY 1381 TACAGGAGACGCTGTGATGAGTGTGCTATCTGACACACTATACATGATTAAGGACATA 1440
Db 1381 TACAGGAGACGCTGTGATGAGTGTGCTATCTGACACACTATACATGATTAAGGACATA 1440
QY 1441 GAATGTGTGCAATGTTTACAGAGTGAAGGCAATGAGGAATGGTGGCTTATGGA 1500
Db 1441 GAATGTGTGCAATGTTTACAGAGTGAAGGCAATGAGGAATGGTGGCTTATGGA 1500
QY 1501 TTTGATTAACAGAGATGATGAGTGGGCTTATGATGGAATTAAGTGTGGATCTT 1560
Db 1501 TTTGATTAACAGAGATGATGAGTGGGCTTATGATGGAATTAAGTGTGGATCTT 1560
QY 1561 GTGGCTGCTTTGGACCCCGTGTCTCTGAGGAGACACTGTGTACGAGACCTTGTGAG 1620
Db 1561 GTGGCTGCTTTGGACCCCGTGTCTCTGAGGAGACACTGTGTACGAGACCTTGTGAG 1620
QY 1621 CATTCGGAAGATTTTTCAGTACAGTTGATGAATTCAGATTTGATAGTTCACAT 1680
Db 1621 CATTCGGAAGATTTTTCAGTACAGTTGATGAATTCAGATTTGATAGTTCACAT 1680
QY 1681 GATGACACATCTCATCTGGGACTTCTTAATGATCCAGCTGCCAAGCTGAACCCGCC 1740

Db	1681	GATGACACAAATCCATCTGTGGACTTCCTAATATATCATCGATGCCAAGCTGAACCCCC	1740
Qy	1741	CGTTCOCCTTCGTGCACATACACCTACATCTCCAGATAAATAAACATACACTGACCTCAT	1800
Db	1741	CGTTCOCCTTCGTGCACATACACCTACATCTCCAGATAAATAAACCATACACTGACCTCAT	1800
Qy	1801	ACTTGCCCGAGAACCCATTAAAGTTGCGGATATTTAAAGTATCTGCGCAATACCAGGATAGC	1860
Db	1801	ACTTGCCCGAGAACCCATTAAAGTTGCGGATATTTAAAGTATCTGCGCAATACCAGGATAGC	1860
Qy	1861	AACAACAGTAAATCAATCAAACTACTGCCCCAGTTTCCCTGAGCTAGCCGAGAGAGAGGGCTT	1920
Db	1861	AACAACAGTAAATCAATCAAACTACTGCCCCAGTTTCCCTGAGCTAGCCGAGAGAGAGGGCTT	1920
Qy	1921	TGAGACCTCCTGTGGGACACAGTTGCTGTGACAGTCGGCCAGAGAGGGTCTACTAGACACA	1980
Db	1921	TGAGACCTCCTGTGGGACACAGTTGCTGTGACAGTCGGCCAGAGAGGGTCTACTAGACACA	1980
Qy	1981	ACTGACTCCTTCAGTGTGCTATCAGAAAGATGTCTTCATCAATTTGTAATGATTGGAAAC	2040
Db	1981	ACTGACTCCTTCAGTGTGCTATCAGAAAGATGTCTTCATCAATTTGTAATGATTGGAAAC	2040
Qy	2041	TTTAAACCTCCCTCGCTCTCTCTCTTTCACCTCTGCACCTAGTTTTCCTCATTTGGTTC	2100
Db	2041	TTTAAACCTCCCTCGCTCTCTCTCTCTTTCACCTCTGCACCTAGTTTTCCTCATTTGGTTC	2100
Qy	2101	CAGACAAGAGTACTTTAATAATATATTTAGTCTTTGGCAGAAAAAATAA	2151
Db	2101	CAGACAAGAGTACTTTAATAATATATTTAGTCTTTGGCAGAAAAAATAA	2151

[illegible]

Db	181	AAATTCACCTTAGACAGCATATCAACAAGCTGTGCAGACCTCGCTGTTAAACCAAGAACAAGTA	240
Qy	241	TGTTTAGCAAGCACTCTCTATGAAAGCACTGAGATTGTGTGGCCAAAACAAACTTGGCCAA	300
Db	241	TGTTTAGCAAGCACTCTCTATGAAAGCACTGAGATTGTGTGGCCAAAACAAACTTGGCCAA	300
Qy	301	GGCACTTCAGATATGATGTGGCCAAAGCAAGGAAACTCTCAGCAAGCTTATGAAAGGAA	360
Db	301	GGCACTTCAGATATGATGTGGCCAAAGCAAGGAAACTCTCAGCAAGCTTATGAAAGGAA	360
Qy	361	AAGAACTGTGTGTCAAAATACCTTGAAGCAGTGTGTGAGATCAGATCAATGGAATTTGTG	420
Db	361	AAGAACTGTGTGTCAAAATACCTTGAAGCAGTGTGTGAGATCAGATCAATGGAATTTGTG	420
Qy	421	GAACATCTTATATCCAAATATGTCTATTAACCAACATGGGCAATTAACCTGTATCTTAA	480
Db	421	GAACATCTTATATCCAAATATGTCTATTAACCAACATGGGCAATTAACCTGTATCTTAA	480
Qy	481	CCATATGTTGAGAGAGATTTTCATMACTGTCTGTCCAGCTCGGGGATTTGATTCATTCGCT	540
Db	481	CCATATGTTGAGAGAGATTTTCATMACTGTCTGTCCAGCTCGGGGATTTGATTCATTCGCT	540
Qy	541	GAGAACATTCGTGTATACCTCGSATATGGCAATACATATGTGTGTGGAACCTTGTGTCAAG	600
Db	541	GAGAACATTCGTGTATACCTCGSATATGGCAATACATATGTGTGTGGAACCTTGTGTCAAG	600
Qy	601	GAATGTACCGAGTACACCTCTGATGGCAGTCTGTGGAAAGACCTTATCGAGAGATGTG	660
Db	601	GAATGTACCGAGTACACCTCTGATGGCAGTCTGTGGAAAGACCTTATCGAGAGATGTG	660
Qy	661	AGGACAGATTTCTGTGTGGAAAGGCTGTGGCAAGAACAGAGATGGGCAAGTATTTATTC	720
Db	661	AGGACAGATTTCTGTGTGGAAAGGCTGTGGCAAGAACAGAGATGGGCAAGTATTTATTC	720
Qy	721	AAAAACAACCTCTGTGAAGGGGAATGTCTCTCCCACTCTTTTATAGACACTTTATCT	780
Db	721	AAAAACAACCTCTGTGAAGGGGAATGTCTCTCCCACTCTTTTATAGACACTTTATCT	780
Qy	781	AAAAATTATCAAGACATTTGAGACATATGAAATCTAATTGAGCATGTGGAAAGCATATGTTA	840
Db	781	AAAAATTATCAAGACATTTGAGACATATGAAATCTAATTGAGCATGTGGAAAGCATATGTTA	840
Qy	841	CAGAGATTCCTACTGCGCAAGTGAACCAAGCAAGAGAGTTACTGTTTACATATGATGAT	900
Db	841	CAGAGATTCCTACTGCGCAAGTGAACCAAGCAAGAGAGTTACTGTTTACAGATATGATGAT	900
Qy	901	CAGAAAAATAGTAAGCGGCTTTCGAGACACACATCAATCAGATCTGGGATPAAAAACATTCG	960
Db	901	CAGAAAAATAGTAAGCGGCTTTCGAGACACACATCAATCAGATCTGGGATPAAAAACATTCG	960
Qy	961	GAATGCAACCGGAATTCACAGGCGCATACAGAGTTCAAGTCTCTGTCTCCAGATATGATGAG	1020
Db	961	GAATGCAACCGGAATTCACAGGCGCATACAGAGTTCAAGTCTCTGTCTCCAGATATGATGAG	1020
Qy	1021	AGAGTGTATCATTAACAGATTCATCGGATTCACAGGTCAGAGTGTGGATGTAAATACAGT	1080
Db	1021	AGAGTGTATCATTAACAGATTCATCGGATTCACAGGTCAGAGTGTGGATGTAAATACAGT	1080
Qy	1081	GAATGTCTTAAACGTTGATTCACCATTTGGAAGCAAGTTCTGCAGCTGTGGCTTCAATAT	1140
Db	1081	GAATGTCTTAAACGTTGATTCACCATTTGGAAGCAAGTTCTGCAGCTGTGGCTTCAATAT	1140
Qy	1141	GGCATGTATGTACCTGTCCAAAGATATGTTCCATTGCTGTATGGATATGGCCCTCCCA	1200
Db	1141	GGCATGTATGTACCTGTCCAAAGATATGTTCCATTGCTGTATGGATATGGCCCTCCCA	1200
Qy	1201	ACTGACATTAACCTTCGGAGGGTGTGTGGGACACCGACCTCTCTCATATGTGTGAAC	1260
Db	1201	ACTGACATTAACCTTCGGAGGGTGTGTGGGACACCGACCTCTCTCATATGTGTGAAC	1260
Qy	1261	TTTGTATGACAAAGTACATTTGTTTCTCATCTGTGGGATGGAACCTTAAAGGTATGAAACACA	1320
Db	1261	TTTGTATGACAAAGTACATTTTCTCTCATCTGTGGGATGGAACCTTAAAGGTATGAAACACA	1320

QY 1321 AGTACTTGATTTGTAAGACCTTAATGACACAAAGAGCATTCCTCTGTCAG 1380
 |||||||
 Db 1321 AGTACTTGATTTGTAAGACCTTAATGACACAAAGAGCATTCCTCTGTCAG 1380
 QY 1381 TACAGGACAGAGGCTGTAGTGTGCTCATCTGACACACTCTGATTTAGGACATA 1440
 |||||||
 Db 1381 TACAGGACAGAGGCTGTAGTGTGCTCATCTGACACACTCTGATTTAGGACATA 1440
 QY 1441 GATGTGTGCTCATCTGATTTAGGACAGTGTAGAGGACATGAGCAATGGTGTGATTCGA 1500
 |||||||
 Db 1441 GATGTGTGCTCATCTGATTTAGGACAGTGTAGAGGACATGAGCAATGGTGTGATTCGA 1500
 QY 1501 TTTGATTAACAGAGAGTATGATCTGAGGCTATGATGAGAAATTAAGTGTGGATCTT 1560
 |||||||
 Db 1501 TTTGATTAACAGAGAGTATGATCTGAGGCTATGATGAGAAATTAAGTGTGGATCTT 1560
 QY 1561 GTGGCTGCTTTGGACCCCGCTGCTCTGACAGGACACTGTCTACAGACCTTGTGAG 1620
 |||||||
 Db 1561 GTGGCTGCTTTGGACCCCGCTGCTCTGACAGGACACTGTCTACAGACCTTGTGAG 1620
 QY 1621 CATTCGGAAGAGTTTGGACATCTGATGATGATTTGATTTGATTTGATTTGATTTGATTT 1680
 |||||||
 Db 1621 CATTCGGAAGAGTTTGGACATCTGATGATGATTTGATTTGATTTGATTTGATTTGATTT 1680
 QY 1681 GATGACACATCTCATCTGAGACTTCTTAATGATCCAGTGCCTCAAGCTGACCCCTCC 1740
 |||||||
 Db 1681 GATGACACATCTCATCTGAGACTTCTTAATGATCCAGTGCCTCAAGCTGACCCCTCC 1740
 QY 1741 CGTTCCTCTCTGACACATCTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1800
 |||||||
 Db 1741 CGTTCCTCTCTGACACATCTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1800
 QY 1801 ACTTGCCGACAGACCTTAATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1860
 |||||||
 Db 1801 ACTTGCCGACAGACCTTAATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1860
 QY 1861 AACCAAGTACATCACTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1920
 |||||||
 Db 1861 AACCAAGTACATCACTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1920
 QY 1921 TGAGACTCTGCTGTTGAGACAGTGTGCTGACAGTGTGCTGACAGTGTGCTGACAGTGTGCT 1980
 |||||||
 Db 1921 TGAGACTCTGCTGTTGAGACAGTGTGCTGACAGTGTGCTGACAGTGTGCTGACAGTGTGCT 1980
 QY 1981 ACTGACTCTGCTGTTGAGACAGTGTGCTGACAGTGTGCTGACAGTGTGCTGACAGTGTGCT 2040
 |||||||
 Db 1981 ACTGACTCTGCTGTTGAGACAGTGTGCTGACAGTGTGCTGACAGTGTGCTGACAGTGTGCT 2040
 QY 2041 TTTTAAACCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 |||||||
 Db 2041 TTTTAAACCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 QY 2101 CAGACAAAGGAGTGTATTAATATATTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2151
 |||||||
 Db 2101 CAGACAAAGGAGTGTATTAATATATTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2151

RESULT 3
 AF129530 2151 bp mRNA linear PRI 31-OCT-1999
 LOCUS AF129530 Homo sapiens chromosome 10 F-box protein Fbw1a (FBW1A) mRNA,
 DEFINITION complete cds.
 ACCESSION AF129530
 VERSION AF129530.1 GI:6164609
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 2151)
 AUTHORS Cenciarelli,C., Chlaue,D.S., Guardavaccaro,D., Parks,W., Vidal,M.
 and Pagano,M.

TITLE Identification of a family of human F-box proteins
 JOURNAL Curr. Biol. 9 (20), 1177-1179 (1999)
 MEDLINE 20003060
 PUBMED 10531035
 REFERENCE 2 (bases 1 to 2151)
 AUTHORS Chlaue,D.S. and Pagano,M.
 TITLE Direct Submission
 JOURNAL Submitted (19-FEB-1999) Pathology, NYU Medical Center, 550 First
 Ave. MSB 548, New York, NY 10016, USA
 FEATURES
 source
 1..2151
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /map="10q24"
 1..2151
 /gene="FBW1A"
 70..1779
 /gene="FBW1A"
 /function="regulates the stability of IkappaB alpha- and
 beta-catenin"
 /note="contains one F-box domain and seven WD40 domains;
 forms an SCF with Skp1 and Cull1"
 /product="F-box protein Fbw1a"
 /protein_id="AA04464.1"
 /db_xref="GI:6164610"
 /translation="MDPAEAVLQEKALFPMNSEREDNNGEPPKIIPEKNSLROTY
 NSCARLCLNDQVCLASTAKTECNVAKRKLANGNSMIVPQRKLSIYEKELCY
 KYEOWESDOVEFVPHLISOMCHQOHGINSYLPMLORDITLAPRGDHLNLT
 LSTLDKSLCAEILVKEKRYTSQDMKLIEMKVRIDSLRLARRGQYLFK
 NRPDGNAPNSFRALYPIKIIQDIETLESNRCGRHSIQRIHCHSETSKGYCYLDY
 DKIVSGLDNNTIKIMDKNTLECKRLIHTGTSVLCLOYDERVITGSSDSVRYWDY
 NTEGMLNLTILHCEAVLILRFNMGMAVCSKDRSLAVMDASPTDITLRYLVGRRA
 VNVVFPDDKYIVSASGDRITKYVNTSTGEFVTLGHRGILACLOYRRLVYSGSSN
 TIRLMDIECGACTRVEGHEHLEVRIRDNRRIVSGADGKIKYWDVLAALDPRPAG
 TILRLTIVHSRVRFLQDFEQLYSSSHDITLIMDFLNDPAAQAEPPNSPRTYTY
 ISR"
 BASE COUNT 628 a 467 c 513 g 543 t
 ORIGIN
 Query Match 100.0%; Score 2151; DB 9; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGGGTGGCTGCGCGCTGCGACCAAGGGCGGCCCGGAGAGAGCCAGTGGCC 60
 |||||||
 Db 1 TGGGTGGCTGCGCGCTGCGACCAAGGGCGGCCCGGAGAGAGCCAGTGGCC 60
 QY 61 TCGGCGATTATGAGACCGGCGCGGCGGTGTCGAAGAAGAGCACTCAAGTTATGAT 120
 |||||||
 Db 61 TCGGCGATTATGAGACCGGCGCGGCGGTGTCGAAGAAGAGCACTCAAGTTATGAT 120
 QY 121 TCCTGAGAGAGAGAGTGTATTAATGAGCAACCCCTAGGAGATATACAGAGAAG 180
 |||||||
 Db 121 TCCTGAGAGAGAGAGTGTATTAATGAGCAACCCCTAGGAGATATACAGAGAAG 180
 QY 181 AATTCATTAGACAGACATACAAAGCTGTGCGACACACTGCTTAAACCAAGAAACAGTA 240
 |||||||
 Db 181 AATTCATTAGACAGACATACAAAGCTGTGCGACACACTGCTTAAACCAAGAAACAGTA 240
 QY 241 TGTTTAGCAAGACAGTGTATGAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
 |||||||
 Db 241 TGTTTAGCAAGACAGTGTATGAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
 QY 301 GGCACCTCCAGATGATGTTGTCGCAAGCAAGAACTCTGACAAAGCTATGAAAGGAA 360
 |||||||
 Db 301 GGCACCTCCAGATGATGTTGTCGCAAGCAAGAACTCTGACAAAGCTATGAAAGGAA 360
 QY 361 AAGGAACCTGTGTGCAAAATACCTTTGAGCAGTGTGTAGAGTGTGAGTGTGAGTGTGAGT 420
 |||||||
 Db 361 AAGGAACCTGTGTGCAAAATACCTTTGAGCAGTGTGTAGAGTGTGAGTGTGAGTGTGAGT 420

QY 421 GAACATCTATATCCAAATGTCATATACCAATGGGACATMAAAGTGTCTTAA 480
 DB 421 GAACATCTATATCCAAATGTCATATACCAATGGGACATMAAAGTGTCTTAA 480
 QY 481 CCTATGTCAGAGAGATTTTCACTGCTGCGACCTGGGGATTGGATCATTCGCT 540
 DB 481 CCTATGTCAGAGAGATTTTCACTGCTGCGACCTGGGGATTGGATCATTCGCT 540
 QY 541 GAGAACATCTGTCATACCTGATGCGCAAAATCACTATGCTGCTGAACCTGGTGCAG 600
 DB 541 GAGAACATCTGTCATACCTGATGCGCAAAATCACTATGCTGCTGAACCTGGTGCAG 600
 QY 601 GAATGTACCGAGTGAACCTGATGAGCATGCTGTGAAGAAAGCTTATCGAGAAATGTC 660
 DB 601 GAATGTACCGAGTGAACCTGATGAGCATGCTGTGAAGAAAGCTTATCGAGAAATGTC 660
 QY 661 AGGACAGATTTCTGTGTGAGAGAGGCTGGCAGAAAGAGAGATGGGACAGATTTATTC 720
 DB 661 AGGACAGATTTCTGTGTGAGAGAGGCTGGCAGAAAGAGAGATGGGACAGATTTATTC 720
 QY 721 AAAAACAACCTCTGAGCGGAAATGCTCCCAACTCTTTTATAGAGCACTTATCCT 780
 DB 721 AAAAACAACCTCTGAGCGGAAATGCTCCCAACTCTTTTATAGAGCACTTATCCT 780
 QY 781 AAAATTTATCAAGACATTTGACAAATAGATCTAATTGAGAGATGTGAAGACATATTTA 840
 DB 781 AAAATTTATCAAGACATTTGACAAATAGATCTAATTGAGAGATGTGAAGACATATTTA 840
 QY 841 CAGAGAAATTCCTGCGGAGTGAAGAAAGAGATTTACTGTTTACAGTATGATGAT 900
 DB 841 CAGAGAAATTCCTGCGGAGTGAAGAAAGAGATTTACTGTTTACAGTATGATGAT 900
 QY 901 CAGAAATAGTAAAGGCGCTTCGAGACAAACAAATCAAGATCTGGATTAACACATTTG 960
 DB 901 CAGAAATAGTAAAGGCGCTTCGAGACAAACAAATCAAGATCTGGATTAACACATTTG 960
 QY 961 GAATGCAAGGAAATTTCTACAGCCCATACAGGTTGCTCTCTCTCCAGTATGATGAG 1020
 DB 961 GAATGCAAGGAAATTTCTACAGCCCATACAGGTTGCTCTCTCTCCAGTATGATGAG 1020
 QY 1021 AGAGTGCATTAACAGATGTCGATGTCGAGGTCAGAGTGGGATGTAAATACAGGT 1080
 DB 1021 AGAGTGCATTAACAGATGTCGATGTCGAGGTCAGAGTGGGATGTAAATACAGGT 1080
 QY 1081 GAAATGCTAAACAGCTTGATTTACCATTTGTAGACAGTTCTGCACTTGCCTTTCAATAT 1140
 DB 1081 GAAATGCTAAACAGCTTGATTTACCATTTGTAGACAGTTCTGCACTTGCCTTTCAATAT 1140
 QY 1141 GGCATGATGTCGACCTGCTCCAAAGATGTTCCATTTGCTGTATGGGATATGGCTCCCA 1200
 DB 1141 GGCATGATGTCGACCTGCTCCAAAGATGTTCCATTTGCTGTATGGGATATGGCTCCCA 1200
 QY 1201 ACTGACATTAACCTGCGAGGAGTGTGTCGAGACACGAGTGGCTGTCATATGTTTACAC 1260
 DB 1201 ACTGACATTAACCTGCGAGGAGTGTGTCGAGACACGAGTGGCTGTCATATGTTTACAC 1260
 QY 1261 TTTGATGACAAATGATTTTCTGCACTGCGGATAGACATTAATGAAGATGAACACA 1320
 DB 1261 TTTGATGACAAATGATTTTCTGCACTGCGGATAGACATTAATGAAGATGAACACA 1320
 QY 1321 AGTACTTGTGAATTTGTAAAGACCTTAATGACAAAGAGATGGCTGTTTACAG 1380
 DB 1321 AGTACTTGTGAATTTGTAAAGACCTTAATGACAAAGAGATGGCTGTTTACAG 1380
 QY 1381 TACAGGACAGGCTGTAGTGTGCTCATTTGACAAACACATGATTAATGGACATA 1440
 DB 1381 TACAGGACAGGCTGTAGTGTGCTCATTTGACAAACACATGATTAATGGACATA 1440
 QY 1441 GAATGTGTCATGTTTAAAGAGTGTGAAGGCAATGAGGAAATGCTGTTGATTTGCA 1500
 DB 1441 GAATGTGTCATGTTTAAAGAGTGTGAAGGCAATGAGGAAATGCTGTTGATTTGCA 1500
 QY 1501 TTTGATTAACAAAGATGTCAGTGGGCTTATGATGAAAAATTAAGTGTGGATCTT 1560

DB 1501 TTTGATTAACAAAGATGTCAGTGGGCTTATGATGAAAAATTAAGTGTGGATCTT 1560
 QY 1561 GTGGCGCTTTGGACCCCGCTGCTCCTGCGAGGACACCTGCTCAAGACCTTTGGAG 1620
 DB 1561 GTGGCGCTTTGGACCCCGCTGCTCCTGCGAGGACACCTGCTCAAGACCTTTGGAG 1620
 QY 1621 CATTCGGAAGATTTTGTGACATAGTTGATGAAATTCAGATTTGTCAGTATGACAT 1680
 DB 1621 CATTCGGAAGATTTTGTGACATAGTTGATGAAATTCAGATTTGTCAGTATGACAT 1680
 QY 1681 GATGACAAATCTCTATCTGCGACTTCTTAATGATCCAGCTGGCCCAAGCTAACCCCC 1740
 DB 1681 GATGACAAATCTCTATCTGCGACTTCTTAATGATCCAGCTGGCCCAAGCTAACCCCC 1740
 QY 1741 CGTTCCCTCTTTCGACATACACCTCATCTCCAGATTAATTAACATACATGACCTCAT 1800
 DB 1741 CGTTCCCTCTTTCGACATACACCTCATCTCCAGATTAATTAACATACATGACCTCAT 1800
 QY 1801 ACTTGGCCAGGACCCATTAAAGTGGGATTTTAACGATCTGCAATACAGATGAGC 1860
 DB 1801 ACTTGGCCAGGACCCATTAAAGTGGGATTTTAACGATCTGCAATACAGATGAGC 1860
 QY 1861 AACACAGTAAACATTAACATTAATGCTGCGAGTTCCCTGAGTACGAGAGAGGCTT 1920
 DB 1861 AACACAGTAAACATTAACATTAATGCTGCGAGTTCCCTGAGTACGAGAGAGGCTT 1920
 QY 1921 TGAGACTCTGTTGGGACACAGTGTGTCAGTGGCGCCAGAGCGTACTACAGACA 1980
 DB 1921 TGAGACTCTGTTGGGACACAGTGTGTCAGTGGCGCCAGAGCGTACTACAGACA 1980
 QY 1981 ACTGACTGCTTCAAGTCTGCTATCAGAAAGATGCTTCTATCAATTTGATGATGGAAC 2040
 DB 1981 ACTGACTGCTTCAAGTCTGCTATCAGAAAGATGCTTCTATCAATTTGATGATGGAAC 2040
 QY 2041 TTTTAAACCTCCCTCCTCCTCCTTTCACCTGACCTGATTTTCCATTTGGTTC 2100
 DB 2041 TTTTAAACCTCCCTCCTCCTCCTTTCACCTGACCTGATTTTCCATTTGGTTC 2100
 QY 2101 CAGACAAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2151
 DB 2101 CAGACAAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2151

RESULT 4
 HSBTRCP
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

HSBTRCP
 Homo sapiens mRNA for beta-transducin repeat containing protein.
 Y14153.1 GI:2995193
 beta-transducin repeats; beta-TRCP gene; WD repeat.
 Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 2151)
 Margolin, F., Hour, S.P., Durand, H., Selig, L., Benichou, S.,
 Richard, V., Thomas, D., Stredel, K., and Benarous, R.
 A novel human WD protein, h-beta trcp, that interacts with HIV-1
 Vpu connects CD4 to the ER degradation pathway through an F-box
 motif.
 Mol. Cell 1 (4), 565-574 (1998)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Mol. Cell 1 (4), 565-574 (1998)
 98325370
 9660940
 2 (bases 1 to 2151)
 Benarous, R.
 Direct Submission
 Submitted (03-JUL-1997) R. Benarous, INSERM - I.C.G.M., Laboratoire
 Interactions Proteiques, CHU Cochin, 24 rue de Fg-St-Jacques, 75014
 Paris, FRANCE
 Location/Qualifiers
 1..2151
 /organism="Homo sapiens"

```
/db_xref="taxon:9606"
/cell_line="Jurkat T-cell"
/cell_type="T lymphocyte"
/tissue_type="lymphoid"
70..1779
/gene="beta-TRCP"
70..1779
/gene="beta-TRCP"
/codon_start=1
/product="beta-transducin repeats containing protein"
/protein_id="CAA4572.1"
/db_xref="GI:2995194"
/translation="MDPAEAVLQEKALFKFNSSREDNGNPEPRKLIPEKNSLROTY
NSCARLCLNOETVCLASTAMKTEVCYKTLKANGTSMVPRKRLISASYEKKEICV
KYFQWSEDOVEVEHLISOMCHYOHGHTISYLKPLMDRPTALPARGLDIANI
LSYDAKSLCAELVCKEYRVSDDLMKILIRVYRTDLSRGLAERNGVOYLFK
NKPPDAPNPNSPTRALYPKTIIDIFETESNMRCGRHSIORSIHSESTGYCLOYD
DKIVSGLRNTIKIMDKNTLECKRLITGHTSVLQYDERVITGSSPSYVWMDV
NGBEMNTLIHCEAVLHFRNNGMYTCKDRSIAVMDASPTDITLRVIGHRAA
VNVVDFDKYIVASGDRITIKVNTSTCEVPTLNGHRIACLOYRDLRVYSSDN
TIRLMDIEGACLRVIEHELYRCIIRPDNRKRIVSGAYDKIKVMDLVVALDPRAG
TLCRLVHSGRVFRQDPDFQIVSSHDDTLILINDLPAAQAPRPSRRTTY
ISR"
repeat_region
259..544
/rpl_family="WD"
/rpl_unit="516..544"
BASE COUNT 628 a 467 c 513 g 543 t
ORIGIN
Query Match 100.0%; Score 2151; DB 9; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCCTGGCTGCGCCCTGGCCACCAAGGGGGGGCCCGCGCGAGAGAGAGCCAGTGGCC 60
Db 1 TGCCTGGCTGCGCCCTGGCCACCAAGGGGGGGCCCGCGCGAGAGAGAGCCAGTGGCC 60
QY 61 TCGGCGATTATGACCCGGCCGAGGGGGGGTGTGCAAGAGAGAGCACTCAAGTTTAAAT 120
Db 61 TCGGCGATTATGACCCGGCCGAGGGGGGGTGTGCAAGAGAGAGCACTCAAGTTTAAAT 120
QY 121 TCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 TCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 AATTCACTTACAGACAGACATCAACAGCTGTGCGAGACTGTGCTTAAACCAAGAAACAGTA 240
Db 181 AATTCACTTACAGACAGACATCAACAGCTGTGCGAGACTGTGCTTAAACCAAGAAACAGTA 240
QY 241 TGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 TGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 GGCACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 301 GGCACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 AAGGAAGCTGTGTCAAAATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 AAGGAAGCTGTGTCAAAATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GAACATCTTATATCCCAATGTGTATACCAACATGGGAGACATAAATCTGATCTTAA 480
Db 421 GAACATCTTATATCCCAATGTGTATACCAACATGGGAGACATAAATCTGATCTTAA 480
QY 481 CCTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 CCTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GAGAAACATTTGTGATACCTGAGATGCCAAATCACTATGTCGCTGAGACTGTGTCAGAG 600
Db 541 GAGAAACATTTGTGATACCTGAGATGCCAAATCACTATGTCGCTGAGACTGTGTCAGAG 600

QY 601 GAATGTTACCGAGTACCTCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 GAATGTTACCGAGTACCTCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 AGAGACATTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 AGAGACATTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 AAAAACAACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 AAAAACAACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 AAAATATACAAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 AAAATATACAAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 CAGAGATTCACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 CAGAGATTCACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 CAGAAATATGTAAGCGGCTTGCAGACAAACACATCAAGATCGGGATAAAAACACATTTG 960
Db 901 CAGAAATATGTAAGCGGCTTGCAGACAAACACATCAAGATCGGGATAAAAACACATTTG 960
QY 961 GAATGCAAGAGATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 GAATGCAAGAGATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 AGAGTATCATTAACAGAGATTCAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 AGAGTATCATTAACAGAGATTCAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 GAAATGCTAAACAGCTGTGATTCACCAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 GAAATGCTAAACAGCTGTGATTCACCAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ACTGACATTAACCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 ACTGACATTAACCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 TTTGATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 TTTGATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 AGTACTTGTGAATTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 AGTACTTGTGAATTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 TACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 TACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 GAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 GAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TTTGATTAACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 TTTGATTAACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 GTGGCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 GTGGCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 CATTCGGAAGAGATTTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CATTCGGAAGAGATTTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1681 GATGACAAATCTCATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
```


BASE COUNT	589 a	528 c	567 g	491 t
ORIGIN	<p> ^{ISr} TLLCIRTLVHSGRVRFLQDFDFQIVSSHDDTILLWDFLNDPAHAAPPSPSKTYIY </p>			

Query Match	78.2%	Score 1682.2	DB 10	Length 2175
Best Local Similarity	88.5%	Pred. NC. 0		
Matches 1909	Conservative	0	Mismatches 228	Indels 20
				Gaps 7

Matches 1909; Conservative 0; Mismatches 228; Indels 20; Gaps 7;

Matches 1909; Conservative 0; Mismatches 228; Indels 20; Gaps 7;

OY	1	TGCGTTGGCTGCGGCTGGACACAAAGGGGGGCCCCGGGAGAGCGAGCCATGGCC	60
Db	22	TGCGTTGGCTGCGGCTGGACAAAGGGGGGGCCCCGGGAGAGCACCCATGATGC	81
OY	61	TGCGGATTTATGAGCCCGGAGGGGCTGCTGCAGAGAAAGGCATCAAGTTTGAAT	120
Db	82	CGGGGATTTATGAGCCCGGAGGGGCTGCTGCAGAGAAAGGCCCTTAAGTTTGAAT	141
OY	121	TGCTCAGAGAGAGAGACTGTATATATGGCGAACCCCTAGAGAAATATATACAGAGAG	180
Db	142	TGCTCAGAGAGAGAGACTGTATATATGGCGAACCCCTAGAGAAATATATACAGAGAG	201
OY	181	AATTCACCTTAGACAGACATACACAGCTGTGTCACAGCTCTGCTTAACCAAGAAACAGTA	240
Db	202	AATTCACCTTAGACAGACTTACACAGCTGTGTCACAGCTTGGATTAACCAAGAACAGTA	261
OY	241	TGTTTAGCAAGAGAGCTATGAAAGATGAGATTTGTTGGCCAAAACAAACTTGGCAAT	300
Db	262	TGTTTAGCAAGAGAGCTATGAAAGATGAGATTTGTTGGCCAAAACAAACTTGGCAAT	321
OY	301	GGCAGCTTCCAGTATGATTTGTGCCAAGCAAGGGAAACTCTCAGCAAGCTATGAAAGGAA	360
Db	322	GGCAGCTTCCAGATGATTTGTGCCAAGCAAGGGAAACTCTCAGCAAGCTATGAAAGGAA	381
OY	361	AAGGAAGCTGTGTCAAACTTTGAGCGAGTGGTCAAGTCAATCAATGGAATTTGTA	420
Db	382	AAGGAAGCTGTGTCAAGTATTTTGAGCGAGTGGTCAAGTCAATGGAATTTGTA	441
OY	421	GAACATCTTATATCCCAATGTGTCTTACCAACATGGGACATAACTCGATCTTAA	480
Db	442	GAACACCTTATATCCCAATGTGTCTACCAACATGGGACATCAACTCTACTTAA	501
OY	481	CCTATGTTCAGAGAGATTTCAATACTGCTGCGCAAGCTCGGGGATTTGATCATATCGCT	540
Db	502	CCTATGTTCAGAGAGATTTCAATACTGCACTGCGCAAGCTCGGGGATTTGATCATATCGCT	561
OY	541	GAGAACATTTCTGTCAATACCTGATGCGCAATCACTATGTGCTGCTGAACCTTGTGTCAAG	600
Db	562	GAGAACATTTCTGTCAATACTTGTGAGCGCAAGTCACTGTGTGCTGCTGAGCTCGTGTCAAG	621
OY	601	GAATGTTACCGAGTGAATCTGTATGGCATGTGTGTGGAAGAGCTTATCGAGAAATGTC	660
Db	622	GAATGTTACCGGCTGTGAGCGCATGTGTGGAAGAGCTTATCGAGAAATGTC	681
OY	661	AGGACAGATTTCTGTGAGAGGCTGTGCGAGAACGAAAGAGTGGGAGACAGTATTTATTC	720
Db	682	AGGACAGACTCTGTGTGAGAGGCTGTGCGAGAACGAGAGGCTGGGAGACAGTATTTATTC	741
OY	721	AAAAACAACCTCTGACGCGAAATGCTCTCCCACTCTTTTATATAGACACTTTATCT	780
Db	742	AAAAACAACCTCTCTATATGAAAGGCTCTCCCACTCTTTTATATAGAGGCTTTATCT	801
OY	781	AAATATATACAAACATTTGAGCAATAGATTTAATGGAGATGTGGAAACATATAGTTTA	840
Db	802	AAATATATACAAACATTTGAGCAATAGATTTAATGGAGATGTGGCAACATATAGTTTA	861
OY	841	CAGAGATTTACTGCGCAATGAAACAAGCAAAAGAGATTACTGTTTACGATATGATAT	900
Db	862	CAGAGATTTACTGCGCAATGAAACAAGTAAAGAGGCTTACTGTTTACGATACGAGAC	921
OY	901	CAGAAATTTACTAAGGGGCTTTCAGAGAACAAACAATCAAGATCTGGGATTAACCAACTTG	960
Db	922	CAGAAATTTACTAAGGGGCTTTCAGAGAACCAACAATCAAGATCTGGGATTAACCAACTTG	981

Qy	961	GAATGCAAGGAAATTCACACAGGCCATACAGGTTGACGTCCTGCTCCAGTAAATGAG	1020
Db	982	GAATGCAAGCGAATTCACAGGGCCACACGGGCTCCGTCCTGCTGCTCAGTACGATGAG	1041
Qy	1021	AGAGTGCATTAACAGAGATCATCGAATTCACAGCGTCACAGTGTGGGATGTTAAATACAGGT	1080
Db	1042	AGGTTGATCATCAAGAGGCTCTCAGACCTCACCGTCACAGATGTGGGATGTTAAATGCGAGGT	1101
Qy	1081	GAATTCGCTAAACACGTTTGATTCACCAATTGTGAAGCATTTCTTGCACTTGCCTTTCATTAAT	1140
Db	1102	GAGATGCTAAACACATTTGATTCACCACTGTGAAGCGCTTGTGCACTGCGCTTCATTAATAT	1161
Qy	1141	GGCATGATGGGACCTCTCCAAAGATTCGTTCCATTTGCTGATAGGAAATAGGCCCTCCCA	1200
Db	1162	GGCATGATGGGACCTCTTCCAAAGACCGTTCCATTCGCTGTGTGGGATATGCTTCCCA	1221
Qy	1201	ACTGCATTTACCCCTCCGAGGGTGTGTCGGACACCGAGCTGCTTCATGTTGTAGAC	1260
Db	1222	ACTGCATCACCCCTCAGAGAGGTGTGTCGGACACCGAGCTGCGGTCAATGTTGTAGAC	1281
Qy	1261	TTTGTATGACAAAGTACATTTGTTTCTGCATCTGGGGATAGAACATTAAGAGTATGGAACACA	1320
Db	1282	TTTGTATGCAAGTACATCGTTTGTGCTCTGGAGATGAGAACCATTAAGAGTATGGAACACA	1341
Qy	1321	AGTACTTGTGAATTTGTAAAGACCTTAAATGAGACAAACAGGACATTGCTTTTGACG	1380
Db	1342	AGTACTTGTGAATTCGTAAAGACCTTAAATGAGACAAACCGTGGCATGCGCTGTTTGACG	1401
Qy	1381	TACAGGACAGGCGTGTATGTAAGTAGTGCCTCATCTGCACACACTATCAGATTATGGACATA	1440
Db	1402	TACAGAGACAGCGTGTGTGTAGAGCGGCTCCTCTGCACACACCTAGGCTGTGGACATA	1461
Qy	1441	GAATCGTGTGATGTTTACAGAGTGTAAAGAAAGCCATGAGAGATTTGGTCGTTGTAAATGCA	1500
Db	1462	GAGTGTGAGATGCTGTGCGAGTGTGTGGAAGGCCATGAGAGATTTGTACGCTGCATTCGA	1521
Qy	1501	TTTGTATTAACAAGAGAGATAGTCAGTGGGGCCTATGATGAGAAAAATTAAAGTGTGGATCTT	1560
Db	1522	TTTGTATTAACAAAAGATATAGTAGCGGAGCCTATGATGAGAAAAATTAAAGTGTGGATCTT	1581
Qy	1561	GTGGCTGCTTTTGACCCCGCTGCTCTGCAGGAGACACTGTCTACGAGACCCCTGTGGAG	1620
Db	1582	ATGGCTGCTTTTGACCGCGCTGCTCAGCGAGGAGCTCTGTCTCGGACACTGTGTGAG	1641
Qy	1621	CATTCGCGAAGATTTTTCGACTACATTTGATGAATTCACATTTGTCAGTATGTTACAT	1680
Db	1642	CATTCGGAAGATTTTTCGCGCTCGAATTTGATGAATTCACAGATTTGTCAATGTTACAT	1701
Qy	1681	GATGACACATCTCTATCTGGGACTTGCCTAAATGATTCAGCTGCCACAGCTGAAACCCCC	1740
Db	1702	GATGACACAAATTCATCTGGGACTTGCCTAAATGATTCAGCTGCCACAGCTGAAACCCCC	1761
Qy	1741	CGTCCCGCTTCGCAACATTAACCTACATCTCCAGATTAATAACATTAACATTAACCTACAT	1800
Db	1762	CGTCCCGCTTCGGAATACACCTACATCTCCAGATTAATAACATTAACATTAACCTACAT	1821
Qy	1801	ACTTGGCCAGACCCATTAAAGTTGCGGTATTTAAAGTATCTGCCAATACAGATGAGC	1860
Db	1822	AATTTGCCAGATTCGTTAATGTTGCAATTTTAAACAGACCTGCCAAGACAGATGAAC	1881
Qy	1861	AACACAGTAAACATCAAACTACTGCCCCAGTTTCCCTG--GACTAGCCGAGAGAGAGGCG	1918
Db	1882	-----AACAAATCAAACTCTACCCGATTTCCCGAGCGGATGAGCAGAGAGAGGCG	1932
Qy	1919	TTTGTAGACGCTGTGGGACACAGTTGGTGCAGCGGGCCAGGACGCTGTACTACACA	1978
Db	1933	TTTGTAGACGCTGTGGGACACAGTGGTGCAGAG--CCGACACAGGAGCGGCTGTCTGCG--	1990
Qy	1979	CAACTGACTGCTTCAAGTCTGCTATCAGAAAGATGCTTCTTATCAATTTGTGAATGTGA	2038
Db	1991	--ACCGGCTGCGCTCAAGTCTGCTATCAGAAAGATGCTT--TATCTGTGTGAATGTATGGA	2047

QY	1372	TGTTGCAGTACAGGGACAGCGCTGGTAGTGGTCATCTGCAACAACATATCAGATTA	1431
Db	1411	TGTTTGGAGTACAGGACAGCGCTGGTAGTGGCTCATCTTGCAACACATATCAGATTA	1430
QY	1432	TGGACATAGAAATGGTGCATGTTTACGAGTGTATTAGAGGCCATGAGGAATGGTGGCT	1491
Db	1471	TGGGACATAGAAATGGTGCATGTTTACGAGTGTATTAGAGGCCCATGAGGAATGGTGGCT	1530
QY	1492	TGTATTCGATTTGATTAACAGAGAGATAGTCAAGTGGGCCCTTATGATGGAAAAATTAAATG	1551
Db	1531	TGTATTCGATTTGATTAACAGAGAGATAGTCAAGTGGGCCCTTATGATGGAAAAATTAAATG	1590
QY	1552	TGGATCTTGGGGTGGCTTTGGACCCCCGTCTCTCGAGGAGACACTGTGTCTACGAGAC	1611
Db	1591	TGGGATCTTGGGGTGGCTTTGGACCCCCGTCTCTCGAGGAGACACTGTGTCTACGAGAC	1650
QY	1612	CTTGTGAGCAATTCGCGAAGAGTTTTCGACTACAGTTTGATGAAATTCAGATTGTCACT	1671
Db	1651	CTTGTGAGCAATTCGCGAAGAGTTTTCGACTACAGTTTGATGAAATTCAGATTGTCACT	1710
QY	1672	AGTTACATGATGACACAACTCTCATCTGGGCACTTCTAAATGATCCAGCTGCCCAAGCT	1731
Db	1711	AGTTACATGATGACACAACTCTCATCTGGGCACTTCTAAATGATCCAGCTGCCCAAGCT	1770
QY	1732	GAACCCCCCGTCCCTCTTCGAACTACACACTACATCTCAGATTA	1779
Db	1771	GAACCCCCCGTCCCTCTTCGAACTACACACTACATCTCAGATTA	1818

QY 724 AACAACTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTATCTTAA 783
 DB 661 AACAACTCTGATATAGAACCTCTCCCAACTCTTTTATAGAGCGCTTATCTTAA 720
 QY 784 ATTATACAGACATTTAGACAAATAGATTAATTTAGATGAGATGGAACATGTTTACAG 843
 DB 721 ATCATACAGACATTTAGACAAATAGATTAATTTAGATGAGATGGAACATGTTTACAG 780
 QY 844 AACATTCATCTGCCGAAGTGAACAAAGCAAGTTTACTGTTTACATGATGATGAG 903
 DB 781 AACATTCATCTGCCGAAGTGAACAAAGCAAGTTTACTGTTTACATGATGATGAG 840
 QY 904 AAAATAGTAAAGGCGCTTGGAGAACAAACATCAGATCTGGGATTAATAACATTTGAA 963
 DB 841 AAGATAGTAAAGGCGCTTGGAGAACAAACATCAGATCTGGGATTAATAACATTTGAA 900
 QY 964 TGCAGAGCAATTTCTACAGAGCCATACAGATCTCTCTCTCTCAGATGATGAGAG 1023
 DB 901 TGCAGAGCAATTTCTACAGAGCCATACAGATCTCTCTCTCTCAGATGATGAGAG 960
 QY 1024 GTGATCATACAGAGATTCAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1083
 DB 961 GTGATCATACAGAGATTCAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
 QY 1084 ATGCTAAACAGCTTATACCATTTGAGACAGATTTGACATTTGACATTAATGAC 1143
 DB 1021 ATGCTAAACAGCTTATACCATTTGAGACAGATTTGACATTTGACATTAATGAC 1080
 QY 1144 ATGATGAGTACCTGCTCCAAAGATCTGTCATTTGCTGATGAGATGAGCTCCCAACT 1203
 DB 1081 ATGATGAGTACCTGCTCCAAAGATCTGTCATTTGCTGATGAGATGAGCTCCCAACT 1140
 QY 1204 GACATTTACCTCCGAGAGGCTGCTGTCGACACGAGTGTGTCATGATGAGCTTT 1263
 DB 1141 GACATTTACCTCCGAGAGGCTGCTGTCGACACGAGTGTGTCATGATGAGCTTT 1200
 QY 1264 GATGACATGATCTTTTCTGATCTGAGGATGAACTATTAAGATGAGACACAACT 1323
 DB 1201 GATGACATGATCTTTTCTGATCTGAGGATGAACTATTAAGATGAGACACAACT 1260
 QY 1324 ACTGTCATTTGATGAGGCTTAAATGAGACAAAGAGGATGCTGCTGTTGCACTAC 1383
 DB 1261 ACTGTCATTTGATGAGGCTTAAATGAGACAAAGAGGATGCTGCTGTTGCACTAC 1320
 QY 1384 AGGACAGGCTGATGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1443
 DB 1321 AGGACAGGCTGATGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1380
 QY 1444 TGTGTCATTTGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1503
 DB 1381 TGTGTCATTTGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1440
 QY 1504 GATGACAGGATGATGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1563
 DB 1441 GATGACAGGATGATGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1500
 QY 1564 GCTGCTTTGAGACCCCTGCTCTGACAGGACACTGCTGTCAGGACCCCTTGGAGCAT 1623
 DB 1501 GCTGCTTTGAGACCCCTGCTCTGACAGGACACTGCTGTCAGGACCCCTTGGAGCAT 1560
 QY 1624 TCCGAGAGATTTTCTGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1683
 DB 1561 TCCGAGAGATTTTCTGATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1620
 QY 1684 GACACATCTCTATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1743
 DB 1621 GACACATCTCTATGAGGCTTAAATGAGACAAAGATGATGAGACATGAA 1680
 QY 1744 TCCGCTTTGAGACCCCTGCTCTGACAGGACACTGCTGTCAGGACCCCTTGGAGCAT 1803
 DB 1681 TCCGCTTTGAGACCCCTGCTCTGACAGGACACTGCTGTCAGGACCCCTTGGAGCAT 1740

QY 1804 TGCCAGGACCCATTAAAGTGGCGTATTTAAGCTATCTGCCAATACCGAGATGAGCAAC 1863
 DB 1741 TGCCAGGACCCATTAAAGTGGCGTATTTAAGCTATCTGCCAATACCGAGATGAGCAAC 1797
 QY 1864 AACGATTAACATCACTACTGCCCCAGTTTCCCTG--GACTAGCCGAGAGGAGGCTTT 1921
 DB 1798 -----AACATCAAACTCTTACCCGATTCGCGAGAGATGAGAGGAGGAGGCTTT 1851
 QY 1922 GAGACTCTGTTGGAGACAGATTTGCTGAGTCCGCGCCGAGAGGCTTACTAGACACAA 1981
 DB 1862 GAGACTCTGTTGGAGACAGATTTGCTGAGTCCGCGCCGAGAGGCTTACTAGACACAA 1907
 QY 1982 CTGACTCTTCACTGCTGCTATCAGAGATGCTTCTATCAATGTAATGATGAGACT 2041
 DB 1908 CCGGCTCTCTACAGTCTGCTATCAGAGATGCTT--TATCTTGTGATGATGAGTGAAC 1966
 QY 2042 TTTAACTCTCCCT 2097
 DB 1967 TTTAACTCTCCCT 2026
 QY 2098 TTTCAAGCAAGGCTGATTTAATATATTTAGTGTGTTTGGCCGAGAAAAA 2151
 DB 2027 TTTCAAGCAAGGCTGATTTAATATATTTAGTGTGTTTGGCCGAGAAAAA 2080

RESULT 9
 AF110396 1979 bp mRNA linear ROD 28-JUN-1999
 LOCUS AF110396
 DEFINITION Mus musculus beta-transducin repeat-containing protein mRNA,
 complete cds.
 ACCESSION AF110396
 VERSION AF110396.1 GI:5230821
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
 1 (bases 1 to 1979)
 AUTHORS Winston, J., Ellledge, S.J. and Harper, J.W.
 TITLE Direct Submission
 JOURNAL Submitted (01-DEC-1998) Biochemistry, Baylor College of Medicine,
 One Baylor Plaza, Houston, TX 77030, USA

FEATURES

source

1..1979
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 1..1710
 /note="beta TRCP"
 /codon_start=1
 /product="beta-transducin repeat-containing protein"
 /protein_id="A041025.1"
 /db_xref="GI:5230822"

CDS

1..1979
 /translation="MDPEAVLQERALKFNNSSREDNNGEPPRKIIPEKNSLRQTY
 NSCARLINDTVCITSTAMKTEVCVAKLANSIMIVPKRLISAVERKEICV
 KYEOWESDDEVEVEHLISQMHYGHINSLKPLQDFTLPALRGLDITENI
 LSYDASISCAELVCKEMRYVSDGMLMKLIRMYRRTDSIRGLAERPMQYLFK
 NKPDENAPNSFVRLYPTKIDIEIESNMCGRSIORIRHCRSTSGVCLQYD
 DOKVLSLRQNTIKIMDKSTLECRILITGTSVLCIDERYITGSSDSIVRWVDY
 NAGMDLTLTHCAVLAHLRNNGMWTCKSDSIADWASPLDITLRLRVIVGHRVA
 VNVVDEDKTIVSASGDRITKVMNTSCFVRPLNGHKGRIACLOYRDLRVVSSDN
 TIRLMDLEGACNLVEGHELVACIRFDRKRVISGRYDGR IKVMDLMAIDPRAPAG
 TILRLTIVHSGRVFRLOFDEFOIVSSHDDTLIMDLNDPNAHAHEPRSPRTTY
 ISR"

BASE COUNT 535 a 480 c 516 g 448 t
 ORIGIN

Query Match 72.9%; Score 1567.2; DB 10; Length 1979;
 Best Local Similarity 88.5%; Pred. No. 0;

Matches 1761; Conservative 0; Mismatches 213; Indels 16; Gaps 5;

QY 70 ATGAGCCCGGCGGAGGCGGCTGCTGCAAGAGGCACTCAAGTTATGAAATTCCTGAG 129
 DB 1 ATGAGCCCGGCGGAGGCGGCTGCTGCAAGAGGCACTCAAGTTATGAAATTCCTGAG 60

QY 130 AGAGAGACTGTAATTAATGGGAAACCCCTAGGAAGATTAATACGAGAGAAATTCACCTT 189
 Db 61 AGAGAGAGCTGTAAATATGGGGAACCCCTAGGAAGATTAATACGAGAGAAATTCACCTT 120
 QY 190 AGACAGACATACACAGCTGTGGCAGACTGCTTAAACCAAGAAACAGATGTTAGCA 249
 Db 121 AGACAGACATACACAGCTGTGGCAGACTGCTTAAACCAAGAAACAGATGTTAGCA 180
 QY 250 AGCAGCTGTATGAACAGTGAATGTTGTGGCCAAACAAACCTGGCAATGGCACTTCC 309
 Db 181 AGCAGCTGTATGAACAGTGAATGTTGTGGCCAAACAAACCTGGCAATGGCACTTCC 240
 QY 310 AGTATGATTTGGTCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 369
 Db 241 AGCAGATGATTTGGTCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 300
 QY 370 TGTGTCAATTAATTAATGAGAGTGTGCAAGTCAATCAAGTGAATTTGTGAGACATCTT 429
 Db 301 TGTGTCAATTAATTAATGAGAGTGTGCAAGTCAATCAAGTGAATTTGTGAGACATCTT 360
 QY 430 ATATCCCAATGTCATTAATCAACATGGGCAATTAATCAATGCTATTAATCAATGTTG 489
 Db 361 ATATCCCAATGTCATTAATCAACATGGGCAATTAATCAATGCTATTAATCAATGTTG 420
 QY 490 CAGAGAGATTTGATTAATGCTGCTGCAAGCTCGGGATTTGATCAATGCTGAGAACAT 549
 Db 421 CAGAGAGATTTGATTAATGCTGCTGCAAGCTCGGGATTTGATCAATGCTGAGAACAT 480
 QY 550 CTGTCTACCTGGAGTCCCAATCAATCAATGCTGCTGCAAGCTGCTGCAAGTATGTTG 609
 Db 481 CTGTCTACCTGGAGTCCCAATCAATCAATGCTGCTGCTGCAAGCTGCTGCAAGTATGTTG 540
 QY 610 CGAGTACCTGTCAGTGAAGTGTGGAAGCAAGCTTATGAGAGAAATGCTGAGACAGAT 669
 Db 541 CGAGTACCTGTCAGTGAAGTGTGGAAGCAAGCTTATGAGAGAAATGCTGAGACAGAT 600
 QY 670 TCTCTGTGAGAGAGCTGCGCAAGCAAGAGATGGGAGCAGTATTTATCAAAAACAA 729
 Db 601 TCTCTGTGAGAGAGCTGCGCAAGCAAGAGATGGGAGCAGTATTTATCAAAAACAA 660
 QY 730 CTTCTGAGAGAGAGTCTCTCCCACTCTTTATAGAGCTTATATCTTAATTAATTA 789
 Db 661 CTTCTGAGAGAGTCTCTCCCACTCTTTATAGAGCTTATATCTTAATTAATTAATTA 720
 QY 790 CAGAGATTTGAGCAATTAATTAATGAGATGTAAGAGACATAGTTACAGAGAAAT 849
 Db 721 CAGAGATTTGAGCAATTAATTAATGAGATGTAAGAGACATAGTTACAGAGAAAT 780
 QY 850 CACTGCCAGAGTGAACCAAGAGATTTACTGTTTACAGTATGATGAGAAATA 909
 Db 781 CACTGCCAGAGTGAACCAAGAGATTTACTGTTTACAGTATGATGAGAAATA 840
 QY 910 GTAGAGGCTTGAAGCAACCAATCAATGATGGGATTAACCAATGATGATGATGATG 969
 Db 841 GTAGAGGCTTGAAGCAACCAATCAATGATGGGATTAACCAATGATGATGATGATG 900
 QY 970 CGAATTTCTCAGAGGCTTACAGGTTAGTCTGCTCTCAGTATGATGAGAGATGATC 1029
 Db 901 CGAATTTCTCAGAGGCTTACAGGTTAGTCTGCTCTCAGTATGATGAGAGATGATC 960
 QY 1030 ATTAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1089
 Db 961 ATTAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 1090 AAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1149
 Db 1021 AAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1150 GTGACCTGCTCCAAAGATGTTTCAATGTTGTATGAGATGATGATGATGATGATGATG 1209
 Db 1081 GTGACCTGCTCCAAAGATGTTTCAATGTTGTATGAGATGATGATGATGATGATGATG 1140

QY 1210 ACCCTCCGAGAGGCTGCTGCTGCAACCCAGCTGCTGCTCAATGTTGTAGACTTTGATGAC 1269
 Db 1141 ACCCTCCGAGAGGCTGCTGCTGCAACCCAGCTGCTGCTCAATGTTGTAGACTTTGATGAC 1200
 QY 1270 AAGTACATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1329
 Db 1201 AAGTACATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 QY 1330 GAATTTGTAAGGACCTTAAATGACACAAACGAGGATGCTGCTGTTTGGCAGTACAGGAC 1389
 Db 1261 GAATTTGTAAGGACCTTAAATGAGGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 QY 1390 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1449
 Db 1321 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
 QY 1450 GCATGTTTACAGAGTGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1509
 Db 1381 GCATGTTTACAGAGTGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1440
 QY 1510 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1569
 Db 1441 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 QY 1570 TTGAGACCCCTGCTGCTGCAAGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1629
 Db 1501 TTGAGACCCCTGCTGCTGCAAGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 QY 1630 AGAGTTTTCAGTACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1689
 Db 1561 AGAGTTTTCAGTACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1620
 QY 1690 ATCTCATCTGGAGCTTCTTAATGATGATGATGATGATGATGATGATGATGATGATG 1749
 Db 1621 ATCTCATCTGGAGCTTCTTAATGATGATGATGATGATGATGATGATGATGATGATG 1680
 QY 1750 TCTGACATACACTTACATGCTCCGATTAATTAACATTAACATTAACATTAACATTA 1809
 Db 1681 TCTGACATACACTTACATGCTCCGATTAATTAACATTAACATTAACATTAACATTA 1740
 QY 1810 GGACCCATTAAAGTTGGGATTTAAAGTATGATGATGATGATGATGATGATGATGATG 1869
 Db 1741 GGACCCATTAAAGTTGGGATTTAAAGTATGATGATGATGATGATGATGATGATGATG 1791
 QY 1870 AACCAATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1927
 Db 1792 AACCAATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1851
 QY 1928 CTTGTTGGAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1987
 Db 1852 CTTGTTGGAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1907
 QY 1988 GCTTCAGTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2047
 Db 1908 GCTTCAGTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1966
 QY 2048 CTTCCCTCTCC 2057
 Db 1967 CTTCCCTCTCC 1976

RESULT 10
 AF099932 1712 bp mRNA linear ROD 21-JAN-1999
 LOCUS AF099932 Mus musculus beta-TCP protein E3RS-Ikappab mRNA, complete cds.
 DEFINITION AF099932
 ACCESSION AF099932.1 GI:4008019
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 1712)

AUTHORS Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A.M.,
Andersen, J.S., Mann, M., Mercurio, F. and Ben-Neriah, Y.
TITLE Identification of the receptor component of the
JOURNAL Ikkappa1alpha-ubiquitin ligase
MEDLINE 99075339
PUBMED 9859996
REFERENCE 2 (bases 1 to 1712)
AUTHORS Yaron, A., Hatzubai, A., Mercurio, F., Manning, A.M., Andersen, J.S.,
Mann, M. and Ben-Neriah, Y.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1998) Immunology, Hebrew University of Jerusalem,
Ein Karem, Jerusalem 91120, Israel
FEATURES
source Location/Qualifiers
1..1712
/organism="Mus musculus"
/db_xref="taxon:10090"
1..1710
/note="beta-transducin repeat-containing protein; receptor
component of Ikba-ubiquitin ligase"
/codon_start=1
/product="beta-TTCP protein E3HS-IkappaB"
/protein_id="AA08701.1"
/db_xref="GI:4008020"
/translation="MDPEAVLYQEKALKFNSSEREDCNNGEPRKTIIPKNSLRQTY
NSCARLCTNDETCVLTSTAKTENCVAKAKLANCTSSMIVPKOKLSASEKEKELCY
KYPWESDQVEVEHLSQMHQGHINSYKPMLODFITLALPRLDIAENI
LSYLDASICAELVCKEWMYKTDGMLIKRMYRTSMGLMGLERGMGQYLFK
NPKPDENAPNSFYALYKPIODIETIESMWRGRSLRORHRSPTSGVCLQYD
DOKIVSGRLDTIKIMDKSTLEKRIITGTGSLCQYBERVITTSSTYRWVWY
NAGEMLNTLHHCAYVHLHRENNMAMTSCSDNSIAVDMASPTDITLRVVLGHRVA
VNVYDFDKTIVASGDRITKVMNTSTCEFRNLNGHRIACIQYDRDLVSGSSDN
TIRYMDIECGACILNLEGHLEIVNCIRPDKRIYSGVADKIKVMDLMAIDLRAPAG
TLCIRTIVESHGRVFRLOPDEFOIVSSSHDITLIMDFLNDPAHAHAPRSPRTYTY
ISR"

BASE COUNT 469 a 399 c 453 g 391 t
ORIGIN

Query Match 67.5%; Score 1451.2; DB 10; Length 1712;
Best local Similarity 90.5%; Pred. No. 0;
Matches 1549; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 70 ATGACCCGGCGGAGGCGGTGCTGCAAGAGAAAGCAAGTTATGATTCCTCGAG 129
DB 1 ATGACCCGGCGGAGGCGGTGCTGCTGAGAGAAAGCGCTTAATGATTCCTCGAG 60
QY 130 AGAAGACCTTAATTAATGCGGAACCCCTAGAGAAATATACAGAGAAATTCATT 189
DB 61 AGAAGACCTTAATTAATGCGGAACCCCTAGAGAAATATACAGAGAAATTCATT 120
QY 190 AGACAGACATCAACAGCTGTGCGAGACTGCTCTTAACCAAGAAACAGATGTTAGCA 249
DB 121 AGACAGACATCAACAGCTGTGCGAGACTGCTCTTAACCAAGAAACAGATGTTAGCA 180
QY 250 AGCACTGCTATGAAGACTGTGAATTTGTGGCCAAACAAACCTGCGCATGGCACTTC 309
DB 181 AGCACTGCTATGAAGACTGTGAATTTGTGGCCAAACAAACCTGCGCATGGCACTTC 240
QY 310 AGTATGATTTGGCCCAAGCAAGCAACCTTCAGCAAGATGAAAAAGAAAGAACTG 369
DB 241 AGCACTGATTTGGCCCAAGCAAGCAACCTTCAGCAAGATGAAAAAGAAAGAACTG 300
QY 370 TGTGTCAAACTTTGAGAGCTGTGAGATGCAATCAATGGAATTTTGGAACATTT 429
DB 301 TGTGTCAAACTTTGAGAGCTGTGAGATGCAATGGAATTTTGGAACATTT 360
QY 430 ATATCCCAATGTTGCTATTACCAACATGGGCACATAACTGATCTTAAACCTATGTTG 489
DB 361 ATATCCCAATGTTGCTATTACCAACATGGGCACATAACTGATCTTAAACCTATGTTG 420
QY 490 CAGAGAGATTTCACTGCTGTCGCCAGCTCGGGGATTTGCATCATATGCGTCGAGAACAT 549
DB 421 CAGAGAGATTTCACTGCTGTCGCCAGCTCGGGGATTTGCATCATATGCGTCGAGAACAT 480

QY 550 CTGTCAATACCTGATGCGCAAAATCACTATGTCCTGCAACTGTGTGCAAGATGTTAC 609
DB 481 CTGTCAATACCTGATGCGCAAAATCACTATGTCCTGCAACTGTGTGCAAGATGTTAC 540
QY 610 CGAGTGAACCTGATGATGCGCATGCTGTGCAAGCTTATGCAAGAAATGTTCAGACAGAT 669
DB 541 CGAGTGAACCTGATGATGCGCATGCTGTGCAAGCTTATGCAAGAAATGTTCAGACAGAT 600
QY 670 TCTGTGTGAGAGAGCGCTGCGCAAGCAAGAGATGGGAGACATTTATTCAAAAACAA 729
DB 601 TCTGTGTGAGAGAGCGCTGCGCAAGCAAGAGATGGGAGACATTTATTCAAAAACAA 660
QY 730 CCTCTGACGGGAATGCTCTCCCACTCTTTTATAGACACTTATATCCTAAATATTA 789
DB 661 CCTCTGATGAGAACGCTCTCCCACTCTTTTATAGACACTTATATCCTAAATATTA 720
QY 790 CAAGACATTTGACATAGAAATCTAATTTGAGATGTTGAGAAAGATGTTTACAGAAAT 849
DB 721 CAAGACATTTGACATAGAAATCTAATTTGAGATGTTGAGAAAGATGTTTACAGAAAT 780
QY 850 CACTGCCGAAGTGAACCAACCAAGAGATTTACTGTTTACAGATATGATGACAGAAAT 909
DB 781 CACTGCCGAAGTGAACCAACCAAGAGATTTACTGTTTACAGATATGATGACAGAAAT 840
QY 910 GTAAAGCGGCTTGGAGACCAACATCAAGATGTGGATTAACAAACATTTGAAATGCAAG 969
DB 841 GTAAAGCGGCTTGGAGACCAACATCAAGATGTGGATTAACAAACATTTGAAATGCAAG 900
QY 970 CGAATTCACAGAGCCATACAGTTCAGTCCCTGCTGCTCAGATATGATGAGAGATG 1029
DB 901 CGAATTCACAGAGCCATACAGTTCAGTCCCTGCTGCTCAGATATGATGAGAGATG 960
QY 1030 ATTAACAGATATCAATGATTCACAGTTCAGTCCCTGCTGCTCAGATATGATGAGATG 1089
DB 961 ATTAACAGATATCAATGATTCACAGTTCAGTCCCTGCTGCTCAGATATGATGAGATG 1020
QY 1090 AACACGTTGATTCACATTTGTAAGCAAGTTCGCACTTCGCTTCAATATGATGATG 1149
DB 1021 AACACGTTGATTCACATTTGTAAGCAAGTTCGCACTTCGCTTCAATATGATGATG 1080
QY 1150 GTGACCTGCTCCAAAGATGCTTCATGCTGTATGATGATGATGATGATGATGATG 1209
DB 1081 GTGACCTGCTCCAAAGATGCTTCATGCTGTATGATGATGATGATGATGATGATG 1140
QY 1210 ACCCTCCGAGGGTGTGCTGCGACACCGAGCTGCTCAATGTTGTAGACTTTGATGAC 1269
DB 1141 ACCCTCCGAGGGTGTGCTGCGACACCGAGCTGCTCAATGTTGTAGACTTTGATGAC 1200
QY 1270 AAGTACATTTGTTGCAATCTGGGATAGAACTTAAGAGTATGAGACAAAGTACTTGT 1329
DB 1201 AAGTACATTTGTTGCAATCTGGGATAGAACTTAAGAGTATGAGACAAAGTACTTGT 1260
QY 1330 GAATTTGTAAGACCTTAATGATGACAAACAGGACATTTGCTGTTTCACTAGAGGAG 1389
DB 1261 GAATTTGTAAGACCTTAATGATGACAAACAGGACATTTGCTGTTTCACTAGAGGAG 1320
QY 1390 AGCTGTGTGATGAGTGGCTCATCTGACAAACATACATATATGAGACATATGATG 1449
DB 1321 AGCTGTGTGATGAGTGGCTCATCTGACAAACATACATATATGAGACATATGATG 1380
QY 1450 GCATTTTACAGATTTAGAAAGCCATGAGAAATTTGCTGTTTCACTAGAGGAG 1509
DB 1381 GCATTTTACAGATTTAGAAAGCCATGAGAAATTTGCTGTTTCACTAGAGGAG 1440
QY 1510 AAGAGATATGATGAGTGGGCTATGATGAAATTAAGTGTGGGATCTTGTGCTGCT 1569
DB 1441 AAGAGATATGATGAGTGGGCTATGATGAAATTAAGTGTGGGATCTTGTGCTGCT 1500
QY 1570 TTGACCCCGCTGCTGCTGCAAGGACATCTGCTAGGAGACCTTGTGAGACATTTCCGGA 1629
DB 1501 TTGACCCCGCTGCTGCTGCAAGGACATCTTGTGCTGAGACATTTGTGAGACATTTCCGGA 1560

Qy	1630	AGAGTTTTCCACATCAAGATTTTGACAAATTCAGATGTCAGAGTTCACATGATGACACA	1689
Db	1561	AGAGTTTTCCCGCCTCCACAGTTGATGTAATTCAGATGTGCAGAGTTCACATGATGACACA	1620
Qy	1690	ATCCCATCATCGGAGCTTCTCTAAATGATCCAGCTGCAGCGCAACCCCGCTGCCCT	1749
Db	1621	ATTCCTCATCTGGAGATTTCTCTGATGATCCAGCTGCTCCACGCTGGAACCCGCCCTCCCT	1680
Qy	1750	TCTCGAACAATCAACCTACATCTCCAGATAAAT	1781
Db	1681	TCTCGACATACACCTACATCTCCAGATAAAT	1712

RESULT	11				
LOCUS	AF112979	1710 bp	mRNA	linear	ROD 02-MAR-1999
DEFINITION	Mus musculus beta-transducin repeat containing protein mRNA,				
	complete cds.				
ACCESSION	AF112979				
VERSION	AF112979.1	GI:4140717			
KEYWORDS					
SOURCE	Mus musculus.				
ORGANISM	Mus musculus.				

REFERENCE	1 (bases 1 to 1710)
AUTHORS	Spencer, E., Jiang, J., and Chen, Z. J.
TITLE	Signal-induced ubiquitination of Ikappaalpha by the F-box protein Slimb/Beta-Trcp
JOURNAL	Genes Dev. 13 (3), 284-294 (1999)
MEDLINE	99145465
PUBMED	9990853
REFERENCE	2 (bases 1 to 1710)

FEATURES	Location/Qualifiers
source	1. 1710
JOURNAL	Submitted (10-DEC-1998) Molecular Biology and Oncology, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-9148, USA
FILE	Direct Submission

CDS

/function= ubiquitinates phosphorylated Ikba"
/note=beta-Trcp; F-box protein; Ikb-ubiquitin ligase;
substrate recognition subunit of SCF complex; similar to
Homo sapiens beta-Trcp and Drosophila melanogaster Slimb"

```
/codon_start=1
/product="beta-transducin repeat containing protein"
/protein_id="AAD04181.1"
/db_xref="GI:4140718"
```

BASE COUNT	469 a	399 c	453 g	389 t
ORIGIN				

Query Match	67.1%	Score 1442.8	DB 10;	Length 1710;
Best Local Similarity	90.2%	Pred. No. 0;		
Matches 1543; Conservative	0;	Mismatches 167;	Indels 0;	Gaps 0

QY	70	ATGACCCCGCGAGGGGCGTGC	TGCAAGAGAAAGCACTCA	GTGTTGTAATTCCTAGAG	129
Db	1	ATGACCCCGCGAGGGGCGTGC	TGCAAGAGAAAGCACTCA	GTGTTGTAATTCCTAGAG	60

Oy	130	AGAAACACTGTATATATATGGCCAGACCCCTAGGAAATATACAGAGAAATTCACCTT	189
Db	61	AGAGAAACACTGTATATATATGGCCAGACCCCTAGGAAATATACAGAGAAATTCACCTT	120
Oy	190	AGACAGACATACAAACAGCTGTGGCAGACTGTCTTAAACCAAGAAACAGTATGTTAGCA	249
Db	121	AGACAGACATACAAACAGCTGTGGCAGACTGTCTTAAACCAAGAAACAGTATGTTAGCA	180
Oy	250	AGCACTGCTATGAAAGCTGAGAAATGTGTGGCCAAAACAAAACCTTGGCATGGCACTGCC	309
Db	181	AGCACTGCTATGAAAGCTGAGAAATGTGTGGCCAAAACAAAACCTTGGCATGGCACTGCC	240
Oy	310	AGATGATTTGTGCCCAAGCAACGGAACACTCAGCAACGCTATGCAAAAAGAAAGAACTT	369
Db	241	AGCATGATTTGTGCCCAAGCAACGSGGAAACCTCAGCAACGCTATGGAAGAAAGAAAGCTT	300
Oy	370	TGTGTCAATTAATCTTTAGCAGCTGTGTAGAGTACATCAAGTGGAAATTTGTGGAAACATCTT	429
Db	301	TGTGTCAAGTAAATTTTGGAGCAGTGTGTAGAGTGTGATCAAGTGGAAATTTGTGGAAACATCTT	360
Oy	430	ATATCCCAATAGTGTCTATTACCAACATGSGGACATTAACCTGTAATCTTAAACCTATGTTG	489
Db	361	ATATCCCAATAGTGTCTACTCTACAGCATGGGCAATCACTCCTACCTAATAAACCTATGCTG	420
Oy	490	CAGACAGATTTCAATTAATCTGCTCTGCCAGCTCGSGGATTTGGATCATATCGCTGAGAACATT	549
Db	421	CAGAGGGATTTCAATTAATCTGCACTGCGCCAGCAGCGGGCTGTGGACACATCGCTGAGAAACATT	480
Oy	550	CTGATATACCTGGAAATGCGCAATCACTATGTCGTGCTGAAACTGTGTGCAAGAAATGGTAC	609
Db	481	CTGTATATCTTTGAGCGCCAACTCACTGTGTGCTGCTGAGACTGTGTGCAAGAAATGGTAC	540
Oy	610	CGAGTACACTTGATGGCATGCTGTGGAAAGACTTATCAGAGAAATGTCAGACAGAT	669
Db	541	CGCGAGACTGGAGCGGCATGCTGTGGAAAAAGCTCATGAGAGATGTCAGACAGCAGC	600
Oy	670	TCCTGTGTGAGAGGCGCTGGCAAGCAAGAGATGGGGCAGTATTTATCAAAAACAAA	729
Db	601	TCCTGTGTGAGAGGCGCTGGCAGAGCGAGAGGCTGGGGCAATACTTATTCAAAAACAAA	660
Oy	730	CCCTCTGAGGGGAAATGCTCTCCCAACTCTTTTATAGACACTTATCTTAAATTTATA	789
Db	661	CCCTCTGATGAGAAACGCTCTCCCAACTCTTTTATAGAGCCCTTATCTTAAATTTATA	720
Oy	790	CAGAATTTGAGACATTAATGAATCTAATGAGATGTGGAAAGCATATGTTACAGAGATT	849
Db	721	CAGAATTTGAGACATTAATGAATCTCAATTTGAGATGTGGGCGCATATGTTACAGAGATT	780
Oy	850	CACCTGCCAGAGGAACAGCAAGCAAGAGATTACTGTTTACAGTATGATGATCAGAAATA	909
Db	781	CACCTGCCAGAGGAACAGCAAGTAAGGGGTTTACTGTTTACAGTATGATGATCAGAAATA	840
Oy	910	GTAAGGGGCTTCCAGACACACACATTAAGATCTGGGATTAATAACACATTTGGAATGCAG	969
Db	841	GTAAGGGGCTTCCAGACACACACATTAAGATCTGGGATTAATAACACATTTGGAATGCAG	900
Oy	970	CGAATTTCCAGAGGCGATACAGGTTACGTCCTCTGTCTCCAGTATGATGAGAGATGATC	1029
Db	901	CGGATTTCTACGGGCGACAGGGGCTCGCTGTGTGTGAGTACAGGTGAGAGGGTATC	960
Oy	1030	ATPAACAGATCATCGGATTCACGCTCAGAGTGTGGATGTAAATPAACAGTAAATGCTA	1089
Db	961	ATPAACAGGCTCTCCAGACTCCACGCTCAGAGTGTGGATGTAAATPAACAGTAAATGCTA	1020
Oy	1090	AACAGCTTAATCACCATTTGGAAGCAAGTTCTGCACTGGGTTCAATTAATGGCATGATG	1149
Db	1021	AACAGCTTAATCACCATCTGGAAGCCGTTCTGCACTGGGTTCAATTAATGGCATGATG	1080
Oy	1150	GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGSGATATGCGCTCCCAACTCAATT	1209
Db	1081	GTGACCTGTTCCAAAGACCGTTCCATGCTGTGTGGATATGCGCTCCCAACTCAATC	1140
Oy	1210	ACCCCTCGGAGGGTGTGTGTGGACACCGGACCTGTCTCAATTTGTAACTTTGATGAC	1269


```

Db 1141 ACCCTAGAGAGGCTGCTGGGAGACACGAGCTGCGCTCAATGTGTATACCTTTGATGAC 1200
OY 1270 AAGTCAATGTTCTTCATCTGGGGATAGACATATAAGGATGACCAAGTACTGT 1329
Db 1201 AAGTCAATGTTCTTCATCTGGGGATAGACATATAAGGATGACCAAGTACTGT 1260
OY 1330 GAATTTGTAAAGACCTTAAATATGACACAAAGGACGATTCGCTGTTTGCATGACGAGGAC 1389
Db 1261 GAATTCGTAAAGACCTTAAATATGACACAAAGGACGATTCGCTGTTTGCATGACGAGGAC 1320
OY 1390 AGGCTGTAGTATGAGGCTTCTGACAAACACTATGATTAATGAGACATAGATGTGT 1449
Db 1321 AGGCTGTAGTATGAGGCTTCTGACAAACACTATGATTAATGAGACATAGATGTGT 1380
OY 1450 GCATGTTTACGAGTGTATGAGGACCTGAGGATTTGGCTGTTTATTCGATTTATAC 1509
Db 1381 GCATGCTCGAGCTGTTGAGAGGCGCATGAGAGGATTTGATGCTGATTCGATTTATAC 1440
OY 1510 AAGAGATATGTCAGTGGGCGCTATGATGAGAAATTAAGTGTGGATCTTGCGCTCT 1569
Db 1441 AAGAGATATGTCAGTGGGCGCTATGATGAGAAATTAAGTGTGGATCTTGCGCTCT 1500
OY 1570 TTGGACCCCGGCTGCTGACGAGGACCTCTGTCGCGAGACCTTGAGAGATTCGGA 1629
Db 1501 TTGGACCCCGGCTGCTGACGAGGACCTCTGTCGCGAGACCTTGAGAGATTCGGA 1560
OY 1630 AAGATTTTTCAGTACAGTTTATGATTCAGATTTGATGATTCAGATTCAGATTCAGAC 1689
Db 1561 AAGATTTTTCAGTACAGTTTATGATTCAGATTTGATGATTCAGATTCAGATTCAGAC 1620
OY 1690 ATTCATATCTGGGACTTCCTTAATATGATTCAGGCGGCAAGTGAACCCCGCTGCTCT 1749
Db 1621 ATTCATATCTGGGACTTCCTTAATATGATTCAGGCGGCAAGTGAACCCCGCTGCTCT 1680
OY 1750 TCTCGACATACCTACATCTCCAGATTA 1779
Db 1681 TCTCGACATACCTACATCTCCAGATTA 1710

```

```

RESULT 12
XELSBTRCP 1671 bp mRNA linear VRT 17-SEP-1993
LOCUS African clawed frog beta-trcp mRNA, complete cds.
DEFINITION M98268
ACCESSION M98268.1 GI:295542
VERSION beta-transducin repeats.
KEYWORDS Xenopus laevis (library: S. cerevisiae expression library of
SOURCE X.laevis oocytes) cDNA to mRNA.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.

```

```

REFERENCE 1 (bases 1 to 1671)
AUTHORS Spevak,W., Keiper,B.D., Stratowa,C. and Castanon,M.J.
TITLE Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
with beta-transducin repeats
JOURNAL Mol. Cell. Biol. 13 (8), 4953-4966 (1993)
MEDLINE 9330289
PUBMED 8393141

```

```

FEATURES
source Location/Qualifiers
1..1671 /organism="Xenopus laevis"
/db_xref="taxon:8355"
/cell_type="oocyte"
/tissue="S. cerevisiae expression library of X.laevis
oocytes"
1..1671 /gene="beta-trcp"
48..1604 /gene="beta-trcp"
/contig_start=1
CDS

```

```

/protein_id="AA02810.1"
/db_xref="GI:295543"
/translation="MEGFSCSLQPTASEREDCNDEPPRKITEKTLRQKLANGT
SMVPRKRLSANVEREKELCYKFEQWSECDQVFEVHLISRMKHOHGHNTYTK
PMLOBDITLPAKGLDHLAENLSTLDKSLASALVCKERTPTSDMLKRLIER
MVRDSILMRILARRRGMQYLFKRPDPDTPPNSTYRALYPRLLIDETIENMCG
RHSIORHCHSETSKGYCLQYDOKIVSGLRNTTRIKINDKNTLRECKRYLMGHSVL
CLQYDERVLIITGSSDSIVRYMDVNTGEMNTLIHCEAVLIHFNMGMAVTCSDRSI
AVMDMAISDITLRLVYVHRAVNVNVEDEDKYIVSASDRTIKVNTSTCEVRILN
GHRGIACIQYRDRLVSSSDNTILFMDIEGACLRVLEHGEHLRCPTRPNKRTYS
GAYDGRKRWMDLVNVALDPNAPATLCLRLVHSESVFRLQYDERQIVSSSHDITLI
MDPLNDPGLA"
polyA_site 1671
BASE COUNT 495 a 341 c 424 g 411 t
ORIGIN
Query Match 47.3%; Score 1017.8; DB 5; Length 1671;
Best Local Similarity 81.8%; Pred. No. 1e-278;
Matches 1175; Conservative 0; Mismatches 262; Indels 0; Gaps 0;
OY 284 AAGCAAACTTGGCCATGGACCTTCAGTATGATTTGCCCCAAGCAAGGAACTCTCAG 343
Db 157 AGACAAACTTGGCCATGGACCTTCAGTATGATTTGCCCCAAGCAAGGAACTCTCAG 216
OY 344 CAAGCTATGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 403
Db 217 CAATTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 276
OY 404 ATCAAGTGAATTTGAGACATCTTATATCCCAATGTTGCTATTCACCAATGGGCAACA 463
Db 277 ATCAAGTGAATTTGAGACATCTTATATCCCAATGTTGCTATTCACCAATGGGCAACA 336
OY 464 TAAACTGTATCTTAACTATGTCAGAGAGATTTAACTGCTGCGCAAGCTCGGG 523
Db 337 TAAACTGTATCTTAACTATGTCAGAGAGATTTAACTGCTGCGCAAGCTCGGG 396
OY 524 GATTTGATTCATTCGCTGAGAAATTCGTCATCTGATGATGATGATGATGATGATGATGAT 583
Db 397 GATTTGATTCATTCGCTGAGAAATTCGTCATCTGATGATGATGATGATGATGATGATGAT 456
OY 584 CTGAACTTGTGTCAGAGAGATGTTACGAGTACCTCTGATGATGATGATGATGATGATGATG 643
Db 457 CAGAACTGTATGATGAGAGTGTATCGAGTACCTCTGATGATGATGATGATGATGATGATG 516
OY 644 TTATGAGAGATGTCAGAGAGATTTCTGTGAGAGGCTGCGCAAGCAAGAGAGAT 703
Db 517 TCATAGAGGAGATGTCGCGAGAGATTTCTTTGAGAGAGGAGAGAGAGAGAGAT 576
OY 704 GGGGACAGTATTTATCAAAAACAAACCTGACGAGGAAATGCTCTCCCAACTCTTTT 763
Db 577 GGGGACAGTATTTATTTAAAAAACAAACCTGACGAGGAAATGCTCTCCCAACTCTTTT 636
OY 764 ATAGAGCATTTATCTTAAATTTATCAAGACATTTAGACATTTAGATTTATTTAGAT 823
Db 637 ACAGAGCGGTTTACCAAAAATTTATCAAGACATTTAGACATTTAGATTTATTTAGAT 696
OY 824 GTGGAAGACATGTTTACAGAGATTTCTGCGGAGGAAAGAAAGCAAGAGAGATTTACT 883
Db 697 GTGGAAGACATGTTTACAGAGATTTCTGCGGAGGAAAGCAAGAGAGATTTACT 756
OY 884 GTTTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 943
Db 757 GTTTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
OY 944 GGGATTAATAAACAATTTGAGAGGCAAGCAATTTGACAGGCGCATACAGTTCAGTCTCT 1003
Db 817 GGGATTAATAAACAATTTGAGAGGCAAGCAATTTGACAGGCGCATACAGTTCAGTCTCT 876
OY 1004 GTTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1063
Db 877 GTTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936

```


RESULT 14	
AB033280	
LOCUS	AB033280
DEFINITION	Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
ACCESSION	AB033280
VERSION	AB033280.1 GI:7209810
KEYWORDS	F-box and WD-repeats protein beta-TRCP2 isoform B.
SOURCE	Homo sapiens fetal lung cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (sites)
AUTHORS	Koike,U., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M., et al.

TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	JOURNAL	FEATURES
Katoh, M.	Molecular cloning and genomic structure of the betaTRCP2 gene on chromosome 5q35.1	Biogen. Biophys. Res. Commun. 269 (1), 103-109 (2000)	2 (bases 1 to 2252)	Katoh, M.	Direct Submission	Submitted (05-OCT-1999) Masaru Katoh, National Cancer Center, Genetics Division, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan (E-mail:mkatoh@nci.go.jp, Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
source	1. .2252	/organism="Homo sapiens"	/db.xref="taxon:9606"	/chromosome="5"	/map="5q35.1"	/tissue_type="Fetal lung"
gene	1. .2252	/gene="BTRCP2"	157. .1746	/gene="BTRCP2"	/codon_start=1	/product="F-box and WD-repeats protein beta-TRCP2 isoform B"
CDS	157. .1746	/gene="BTRCP2"	/codon_start=1	/product="F-box and WD-repeats protein beta-TRCP2 isoform B"	/protein_id="BAA92330.1"	/db.xref="GI:7209811"
		/translation="MEPDSVIEDKTIELMTSVMEQNEDESKKTLTQISNGTSSIVSRKPSSEGNOKEDKICIKYEPOMSENOVEVEHLISRMCHTOHINSYELKPMJDFETALPEQGLDIIANIITSIYDASLSIAELVCEKQWYISBEMLRKLIEMVYTRDPLWGLSERGRMPQYLFKNRPIDGPNFSFSLPKYIODEIIESNMGRGRLNDRIOCSSEKSGYVCIQYDDEK1ISGLDMSN1KIMDTSLECLVLTGHSVLCILOYERIVITVSSDSIVRWVDNTEVEINFLIHNEAVLH1RFSNGIAMVTCSDRS IAWMDASDTIDLERVYLVGHRAAVNVDEDDK1IVASGSDRI1VWMS7NCEPRLNGKRCIKILOYRDLVYVSGSSDNTIRLMDIECGAC1LVLEBHELVNCEIPFDRK1VSGAYDCIKIWMDLQALDIPRAPASTLC1RLTVLHSHSRVYR1RDPEFOI1SSHDOT1ILMDPLNVPSQVNERBSRSRYTYISR"				
BASE COUNT	598 a	501 c	558 g	595 t		
ORIGIN						
Query Match	41.1%	Score 883.4	DB 9	Length 2252		
Best Local Similarity	74.8%	Pred. No. 2.3e-240				
Matches 1122	Conservative	0	Mismatches 371	Indels	6	Gaps 1
QY	298	AATGACCTCCAGTATGTGATTTGTGCCCAAGCAAGGAAACTCTGACGACGTATGAAAAG	357			
Db	271	AATGGAAACATCATCTGTGATCGTCCAGAAAGAGGCCATCGAAGGAAACTGTCAAAA	330			
QY	358	GAAGAAGACGTGTGTCAATACTTTAGCAGTGTGTCAGAGTCAGATCAAGTGGAAATT	417			
Db	331	GAAGAAACACTGTGTATTAATATATTTTGACCAAGTGTCTGAATCAATCAAGTGGAAATT	390			
QY	418	GTGGACATCTTATATCCCAATGTGTATTTACCAATGAGGACATTAATCTGATCTT	477			
Db	391	GTGGACATCTTATTTTCAGAAATGTGTATTATACGATGGAATTTAACTCTTACCTG	450			
QY	478	AAACCTATGTCAGAGAGATTTCATCTCTGTCGACAGCTCGGGGATTTGGATCAATTC	537			
Db	451	AAAGCATGTGTCAGAGGGGACTTTATTCGCCCTTTACCAAGCAAGGCTTAATATCAATA	510			
QY	538	GGTGAACAATCTGTGATCTTACCTGGATGCCAAATCACTATGTGCTGTAACCTGTGTGC	597			
Db	511	GGCAAAACATCTTTTCGTACTGTGATGAGCCAGGTCTCTGTGTGTCACCAAGCTGTATGT	570			
QY	598	AAAGATGTTACCGATCACTCTGATGGCATGGTGTGGGAAGGATTTCAGAGAAATG	657			
Db	571	AAAGATGTCACGAGTATCTCAGAAAGAAATGCTTTTGGAAAGAGCTGATTTAACGAATG	630			
QY	658	GTCAGACAGATTTCTCTGTGAGAGAGCCCTGGCAAGACGAAGAGATGGGACAGTATTTA	717			
Db	631	GTCAGACAGATCCCTTTCGTGAAAGGACTTTGAAAGAAAGAGGAGGATGAGTACTG	690			

Oy	718	TTCAAAACAAACCTCCCTGACGGGAATGCTCTCCCAACTCTTTTATACAGCACTTAT	777
Dp	691	TTTTAAAACACAGCCACAGATGCG-----CCCTCAATTCATTATTTATAGCTATTATAC	744
Oy	778	CCTAAATTTATACAGACATTTGACACATAGATCTAATTTGAGATGTGGAACATAGT	837
Dp	745	CCAAAGATTATCCAGGATATTTAGAACCTATATGAATCTTAAGTGGCGGTGTGGACACACAC	804
Oy	838	TTACAGACAATTTCACTGGCCGAAGGAAACAAAGCAAGAGTTTAACTGTTACAGTATGAT	897
Dp	805	TTTGCAGAGATTTCAAGTGGCGCTCTGAAAATATATGAAGGTGCTACGTTTACAGTACGAT	864
Oy	898	GATCAAAAATAGTAAAGCGCCTTCCGAGACAAACCAATCAAGATCTGCGATTAACAAACA	957
Dp	865	GATGAAAAAATTTATCAGTGGCGCTACAGAGATATTTCTATTAAGATATGGATTAACACCAG	924
Oy	958	TTTGAATGCAAGCCGAATTTCTACAGGCCATACAGGTTTCACTCCTCTGTCTCCAGTATGAT	101
Dp	925	CTGGAAATTTTGAAGTGTATTAACAGGACACACAGGCTCTGCTCTGTCTGCAAGTATGAT	984
Oy	1018	GAGAGAGTATCATACAGAGATCATCGGATTCGATCCAGGCTCAGAGTGGGGATGTAAATATAC	107
Dp	985	GAGCTGTTCATTTGTAAGTGGCTCTTCACGATTTCTACGTTGAGATGTGAGATGTGAACAG	104
Oy	1078	GGTGAATGCTAAACAGGTTGATTCACCATTTGTGAMACAGTTCTGCACTTGCTTTCAAT	113
Dp	1045	GGTGAAGTTTAAACATTTATCTACACCAACATAGAGCTGTATTTGCACCTACGCTTCAG	110
Oy	1138	AATGGCAGATGGTGGACCTGCTCCAAAGATTCGTTCCATTCGCTGATATGGATATGCGCTCC	119
Dp	1105	AATGGACATGATGGTGAACCTGTTCCAAAGACCGCTCCATTTGCTGTGGGACAGAGGCTTCT	116
Oy	1198	CCAACTGCATTTACCCCTCCGAGCGGTCTGTGGTGGAGACACGACCTCTGTCAATGTTGTA	125
Dp	1165	GGCAGCCGACATCACCTTTACGCCGCTGTCTGTTGGCCACCGGGCTGGCCTCATGTATGTA	122
Oy	1258	GACTTTGATGCAAGTACATTTGTTTCGCACTCTGGGGATATGAACCTATAAGGTATGAAAC	131
Dp	1225	GACTTTGACGCAAGTACATTCGTCTGCTGCTGTGACAGGACACATCAAAAGTCTGGACG	128
Oy	1318	ACAAGTACTTGTGAATTTGTAAGACCTTTAAATGGACAAACAGAGGACATTTGGCTTTTG	137
Dp	1285	ACGAGCACACTGTGAATTTGTTCTGACTCTAAATGGGACAAAGCGGGCATTTGCTCTGTCTC	134
Oy	1378	CAGTACAGGACAGGCTGTGATAGTAGTGCGTCACTGTGCAACACTATACGATATATGGAC	143
Dp	1345	CAGTACAGGGGATCCCGCTGGTTGTAGTGATCATACGATTAATCCATTTAAGGCTCTGGAT	140
Oy	1438	ATAGAAATGGTGTGATTTTAAACGAGTGTTAGAAGCGCATAGGAATTTGGTGGCTGTATT	149
Dp	1405	ATTGAATGTGTGCTCTGTTAAGAGTCTTAAGAGGACATGAAGAATTTGGTCCATGCAATC	146
Oy	1498	CGATTTGATTAACAAGAGGATATAGTACGTGGGCGCTATGATGTGAAGAAATTTAAAGTGGGAT	155
Dp	1465	CGGTTTGTATTAACAAGAGGATTTGTCAGTGGGGCTATATGATGGAAAAATTTAAAGTTTGGAC	152
Oy	1558	CTTGTGGCTGCTTTTGACCCCGCTGCTCTGCAAGGACACACTGTGTACGGACCCCTTGTG	161
Dp	1525	TTTGCAAGCTGTCTTGACCCCTCGAGGCCACAGCAAGACATTTGTTTGGCACAATGGTG	158
Oy	1618	GAGCATTTCCGGAAGATTTTTCGACTACATTTGATGAATTCAGATTTGCAAGTATGTA	167
Dp	1585	GAAATTTCTGACGCTGTGTTCGGCTCCAGTTTGATGATGATTTTCCATTCATCACACCTCC	164
Oy	1678	CATGATGACACAAATCTGATCTGGGACTTGCCTAAATGATACAGCTCCCAAGCTGAACCC	173
Dp	1645	CATGATGACACTATTTGATTTGGGATTTCTTAATATGTGCTCCAGTGGCCAGAAATGAG	170
Oy	1738	CCCGGTTTCCCTTCTGCAACATACACTACATATCTCAGATTAATTAACCATACACTGAC	1796
Dp	1705	ACCGGTTTCTCCCTCCACAAATACACTTCAATCTCAGATTAACAAAGTCTCCACTGTTCAC	1763

FEATURES	source
LOCUS	AB033281
DEFINITION	Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein isoform C, complete cds.
ACCESSION	AB033281
VERSION	AB033281.1
KEYWORDS	BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform C.
SOURCE	Homo sapiens fetal lung cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Kolke,J., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M. and Katoh,M.
TITLE	Molecular cloning and genomic structure of the betaTRCP2 gene on chromosome 5q35.1
JOURNAL	Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
MEDLINE	20160458
REFERENCE	2 (bases 1 to 2274)
AUTHORS	Katoh,M.
TITLE	Direct Submission
JOURNAL	Submitted (05-OCT-1999) Masaru Katoh, National Cancer Center, Genetics Division, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan (E-mail:match@nc.c.go.jp, Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES	Location/Qualifiers
source	1..2274
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="5"
	/map="5q35.1"
	/tissue="type="fetal lung"
	1..2274
gene	/gene="BTRCP2"
	140..1768
CDS	/gene="BTRCP2"
	/codon_start=1
	/product="F-box and WD-repeats protein beta-TRCP2 isoform C"
	/protein_id="BAA92331.1"
	/db_xref="GI:7209813"
	/translation="MEPDSVIEDTILMCSVPSRIIGCANIVESMCALSCDMSPSVRQIOISNGTSIVSRKRPSPEQNYQKEDLCITYFDWSSDVEVEVHLISMCHTHOHNINSLKPLQRFDTLALPEQGLHIAINILSYDARSICAELVCKEMQVISEGLMKKILERNVRDPLMKGLSERGMQVLFKNRPDGPNSFYSRLYPIIDIDIEFTISMWRCGRNHLQRIQCRSENSKGVYCIQYDDEKILISGLRNSIKIMDKTLELKVITYGHSVYACIQYDERIVYTGSSDSTVRVQNTGEVINTLHHNEAVYLHFEFSGLVWYTCSDRSIAVWDMASVDITLRVLVGHRAVNVVDDPKYIVASGDRITKWNSTCEAFRLTNGHGRKIGACIQYRDLVYSSSDMTILMDIEGACIRVLEGHELYRCIRFDNKRIVYSAGDKIKYKWDLOALDPPRAPASTLCRLTLVHSGVFRFLQDFEQLISSSHDPTILMDLNPSPSAQNFTSPSRITYYSIR"
BASE COUNT	587 a 514 c 566 g 607 t
ORIGIN	
Query Match	41.1%; Score 883.4; DB:9; Length 2274;
Best Local Similarity	74.8%; Pred. No.2.3e-240;
Matches 1122;	Conservative 0; Mismatches 371; Indels 6; Gaps 1;
Y	298 AATGCACTTCAGATGATGTTGTGCCAAGCAAGCAAACTCTCAGCAAGCATGTAAG 357
Db	293 AATGGAATCATCATGTGATGCTGCTCCGAAAGAGGCGCATGAGAAAGCAATCATCAAAA 352
Y	358 GAAAGGAACGTGTGCGAAATACCTTTAGACAGAGTGTCTAGAGTCAATCAAGTGAATTT 417
Db	353 GAAAGACCTGTGATTAATTAATTTTACCAAGTGTGTGATCGAGTCAAGTGAATTT 412
Y	418 GTGGAACATCTTATATCCCAATGTGTGATTACCAACATGGGACATTAACCTCGATCTT 477
Db	413 GTGGAACATCTTATTTTACGAAATGTGTGATTAATTCAGCATTAATTAACCTTACTG 472
Y	478 AACTGATGTCAGAGAGATTTCATTAACCTCTGCCAGCTGGGGATTGGATCATATC 537

Db 473 AAGCCATGTCGACGGGACTTTATTACGCTTTACAGACAGGCTTAGATCATA 532
QY 538 GCTGAGACATTTCTGTCTACCTGTAGATGCCAAATTCACATATGCTGCTGACCTTGTGTC 597
Db 533 GCAGAAAAACATTTCTGTCTACCTGTAGATGCCAGGCTGCTGTGTCACAGACGCTGTATGT 592
QY 598 AAGGATGTCAGGAGTGTGCTGTATGCTGTGTGGAAGAGCTTTACAGAGAAAG 657
Db 593 AAGGATGTCAGGAGTGTGCTGTATGCTGTGTGGAAGAGCTTTACAGAGAAAG 652
QY 658 GTCAGACACATTTCTGTGTCAGGAGCTGTGACAGAGAGAGATGGGACATTTTA 717
Db 653 GTACGACATATCTCCCTATGTGGAAGAGCTTTGGAAGAGAGAGGATGATGCTG 712
QY 718 TTTAAAAACAACCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACATTTAT 777
Db 713 TTTAAAAACAACCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGATTTATAC 766
QY 778 CTTAAATTTACAGACATTTGACATATGATGATGATGATGATGATGATGATGAT 837
Db 767 CCAAGATTTATCCAGATATAGAGACTATGAAATCTTACCTGCGGTGTGACACACAC 826
QY 838 TTTACAGAGATTTCTGCTGACGAGTGAAGAGAGAGTGTACGTTTACAGATGAT 897
Db 827 TTGAGAGAGATTTCTGCTGACGAGTGAAGAGAGAGTGTACGTTTACAGATGAT 886
QY 898 GATCAGAAAAATTTAGTACGCGGCTTTGAGACACACATCAAGATCTGGATTTAAACACA 957
Db 887 GATCAGAAAAATTTATCAGTGGCTTACAGATTAATTTATTAAGATTAAGATTAAGATTAAG 946
QY 958 TTGGAATGTCAGGAAATTTCTCAGAGGCTTACAGGCTTCTGCTGCTGCTGCTGCTGCT 1017
Db 947 CTGGAATGTTGAAAGTGTAAAGGTTTAAAGGAGACACAGAGCTCTGCTGCTGCTGCTGCT 1006
QY 1018 GAGAGAGTGTATCATACAGAGATTCAGGATTCAGGATTCAGGATTCAGGATTCAGGATTCAG 1077
Db 1007 GAGGCTGTCTATTTAGTGTCTCTCAGATTTCTGAGATTCAGGATTCAGGATTCAGGATTCAG 1066
QY 1078 GGTGAATGCTTAAACAGGATTTCAACGATTTGAAAGAGTTCGACATTTGCGTTCAAT 1137
Db 1067 GGTGAATGCTTAAACAGGATTTCAACGATTTGAAAGAGTTCGACATTTGCGTTCAAT 1126
QY 1138 AATGGATGATGTCAGCTGCTCCAAAGATGTCCTATGCTGTATGATGATGATGATGATGATGAT 1197
Db 1127 AATGGATGATGTCAGCTGCTCCAAAGATGTCCTATGCTGTATGATGATGATGATGATGATGAT 1186
QY 1198 CCAACTGACATTTACCTCCGAGGCTGTGTGTGAGACACGAGCTGCTGCTGCTGCTGCTGCTGCT 1257
Db 1187 GCGACCGACATTTACCTCCGAGGCTGTGTGTGAGACACGAGCTGCTGCTGCTGCTGCTGCTGCT 1246
QY 1258 GACTTGTATGACAGTATGCTTGTGTCATCTGAGTGTGAGATTTAAAGGATTTGAGAT 1317
Db 1247 GACTTGTATGACAGTATGCTTGTGTCATCTGAGTGTGAGATTTAAAGGATTTGAGAT 1306
QY 1318 ACAAGTACTTGTGAATTTGTAAGGACCTTAATGACACAAACGAGGATTTGCTGTTTG 1377
Db 1307 ACAGACACCTGTGAATTTGTCGACTCTCAATGAGGACAAAGGAGGATTTGCTGCTGCTGCTGCT 1366
QY 1378 CAGTACAGGAGACAGCTGT 1437
Db 1367 CAGTACAGGAGACAGCTGT 1426
QY 1438 ATAGAAATGTCGTCATTTACAGTGTGTTAGAGGACATGAGGATTTGTCGCTGTCATTT 1497
Db 1427 ATAGAAATGTCGTCATTTACAGTGTGTTAGAGGACATGAGGATTTGTCGCTGTCATTT 1486
QY 1498 CGATTTATTAACAAGAGATAGT 1557
Db 1487 CGGTTTATTAACAAGAGATAGT 1546
QY 1558 CTTGTGCTGCTTTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1617
Db 1547 TTGCAAGCTGCTTTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1606

QY 1618 GAGCATTCGGAGAGATTTTTCAGCTACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1677
Db 1607 GAACTTCTGACGCTGT 1666
QY 1678 CATGATGACAAATTCATGCTGT 1737
Db 1667 CATGATGACAAATTCATGCTGT 1726
QY 1738 CCCCCTTCCCTTCTGCAACATACACCTTACATCTCCAGATTAATTAACATTAACATTAACATTAACATTAACAT 1796
Db 1727 ACCGCTTCTCCCTCCAGAAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1785

Search completed: February 21, 2003, 23:50:02
Job time : 5571 secs